

STIC-Biotech/ChemLib

124100

From: Schnizer, Holly  
Sent: Tuesday, June 08, 2004 12:11 PM  
To: STIC-Biotech/ChemLib  
Subject: seq. search in appl. no. 09/509,994

Please search all databases for the following sequences:

aa 19-516 of SEQ ID NO:1  
aa 19-516 of SEQ ID NO:2

Thank you.

Holly Schnizer  
AU 1653  
Office: REM 3C79  
Mailbox: REM 3C70  
(571)272-0958

CRIE

SUMMARY OF SEQ SRCH  
PLEASE ~~keep~~ <sup>SCAN</sup> w/ SEARCH RESULTS

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 6/9/04  
Date Completed: 6/9/04  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: 2  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: ASP  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:39:53 ; Search time 53.5 Seconds  
(without alignments)  
2630.068 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516  
Perfect score: 2830  
Sequence: 1 APAEPQGGSCQVHDCFAAL.....PSTPGSTLTPPAVLVHSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |       |         |     |        |          |           |
|------------|-------|---------|-----|--------|----------|-----------|
| Result No. | Score | Query % |     | Length | DB       | ID        |
|            |       | Match   |     |        |          |           |
| 1          | 2830  | 100.0   | 498 | 2      | AAR84185 | Human der |
| 2          | 2830  | 100.0   | 498 | 2      | AAW01600 | Thrombomo |
| 3          | 2830  | 100.0   | 516 | 2      | AAR22016 | Truncated |
| 4          | 2830  | 100.0   | 516 | 2      | AAY09347 | Human thr |
| 5          | 2830  | 100.0   | 516 | 3      | AAV83934 | Human thr |
| 6          | 2830  | 100.0   | 516 | 3      | AAV69529 | Human thr |
| 7          | 2830  | 100.0   | 516 | 6      | ABU08701 | Thrombomo |
| 8          | 2830  | 100.0   | 516 | 7      | ADA37271 | Human sol |
| 9          | 2830  | 100.0   | 575 | 2      | AAR11534 | Human thr |
| 10         | 2830  | 100.0   | 575 | 2      | AAR41806 | Thrombomo |
| 11         | 2830  | 100.0   | 575 | 2      | AAR43031 | Human thr |
| 12         | 2830  | 100.0   | 575 | 5      | AAR17521 | Human ful |
| 13         | 2830  | 100.0   | 575 | 5      | AAR23026 | Human thr |
| 14         | 2827  | 99.9    | 557 | 5      | AAE17528 | Human thr |
| 15         | 2827  | 99.9    | 557 | 5      | AAE17525 | Human thr |
| 16         | 2827  | 99.9    | 557 | 5      | AAE17525 | Human thr |
| 17         | 2827  | 99.9    | 557 | 5      | AAE23031 | Human thr |
| 18         | 2826  | 99.9    | 557 | 5      | AAE23028 | Human thr |
| 19         | 2826  | 99.9    | 516 | 2      | AAE22018 | Human thr |
| 20         | 2826  | 99.9    | 516 | 3      | AAV09348 | Human thr |
| 21         | 2826  | 99.9    | 516 | 3      | AAV83935 | Human thr |
| 22         | 2826  | 99.9    | 516 | 6      | ABU08703 | Thrombomo |
| 23         | 2826  | 99.9    | 516 | 7      | ADA37275 | Human sol |
| 24         | 2826  | 99.9    | 557 | 5      | AAE17523 | Human thr |
| 25         | 2826  | 99.9    | 575 | 6      | ABG72575 | Human thr |
| 25         | 2826  | 99.9    | 575 | 7      | ABM78950 | Breast ca |

|    |      |      |     |   |                    |
|----|------|------|-----|---|--------------------|
| 26 | 2826 | 99.9 | 575 | 7 | ADBS4760 Human Pro |
| 27 | 2826 | 99.9 | 575 | 7 | ADD48168 Human Pro |
| 28 | 2826 | 99.9 | 631 | 6 | ABG72572 Human thr |
| 29 | 2825 | 99.8 | 498 | 3 | AAY67402 Novel sug |
| 30 | 2825 | 99.8 | 557 | 5 | AAR17596 Human thr |
| 31 | 2824 | 99.8 | 497 | 2 | AAR69520 rSTM prot |
| 32 | 2824 | 99.8 | 557 | 5 | AAR17530 Human thr |
| 33 | 2824 | 99.8 | 557 | 5 | AAR17598 Human thr |
| 34 | 2823 | 99.8 | 557 | 5 | AAR17526 Human thr |
| 35 | 2823 | 99.8 | 557 | 5 | AAR23029 Human thr |
| 36 | 2822 | 99.7 | 516 | 2 | AAR22017 Human thr |
| 37 | 2822 | 99.7 | 557 | 5 | AAR17593 Human thr |
| 38 | 2822 | 99.7 | 557 | 5 | AAR17595 Human thr |
| 39 | 2822 | 99.7 | 557 | 5 | AAR17597 Human thr |
| 40 | 2822 | 99.7 | 557 | 5 | AAR17529 Human thr |
| 41 | 2822 | 99.7 | 557 | 5 | AAR17527 Human thr |
| 42 | 2822 | 99.7 | 557 | 5 | AAR23030 Human thr |
| 43 | 2821 | 99.7 | 557 | 5 | AAR17592 Human thr |
| 44 | 2821 | 99.7 | 557 | 5 | AAR17594 Human thr |
| 45 | 2821 | 99.7 | 557 | 5 | AAR17591 Human thr |

ALIGNMENTS

RESULT 1  
AAR84185  
ID AAR84185 standard; protein; 498 AA.  
XX AAR84185;  
XX  
DT 05-JUN-1996 (first entry)  
XX  
DE Human derived thrombomodulin.  
XX  
KW Human thrombomodulin; hepatopathy; fulminant hepatitis.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Protein 349.463  
FT /note= "fragment AAR84184"  
XX  
PN WO9528953-A1.  
XX  
PD 02-NOV-1995.  
XX  
PF 10-APR-1995; 95WO-JP000704.  
XX  
PR 20-APR-1994; 94JP-00081196.  
XX  
PA (ASAH ) ASAH KASEI KOGYO KK.  
XX  
PI Fujiwara K, Mochida S;  
XX  
DR WPI; 1995-382844/49.  
XX  
PT Hepatopathy treatment pref. comprising human derived thrombomodulin -  
XX useful against, e.g, fulminant hepatitis.  
XX  
PS Claim 2; Page 17-20; 29pp; Japanese.  
XX  
CC The human derived thrombomodulin AAR84185, and its fragment AAR84184 can  
XX be used in the treatment of hepatopathy, including fulminant hepatitis  
XX  
SQ Sequence 498 AA;

Query Match 100.0%; Score 2830; DB 2; Length 498;  
Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 APAEPQGGSCQVHDCFAALPGPATFLNASQICDGLRGLMTVRSSVAADVISLLNGD 60  
|||||

DB 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 QY 61 GGVGRRRLWIGLQPLPGCGDKRLGRLGPGFOWTGDNNTSYSRWARLDLNGAPLCGLPLCV 120  
 DB 61 GGVGRRRLWIGLQPLPGCGDKRLGRLGPGFOWTGDNNTSYSRWARLDLNGAPLCGLPLCV 120  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGMAAAAVSITYGTPFA 180  
 DB 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGMAAAAVSITYGTPFA 180  
 QY 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 DB 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMCTGYRLAA 300  
 DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMCTGYRLAA 300  
 QY 301 DOHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
 DB 301 DOHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
 QY 361 QPLNQTSLYLCVCAEGEAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPBGYILDDGFI 420  
 DB 361 QPLNQTSLYLCVCAEGEAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPBGYILDDGFI 420  
 QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGSEPPPS 480  
 DB 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGSEPPPS 480  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 481 PTPGSTLTTPPAVGLVHSG 498

RESULT 2  
 AAW01600  
 ID AAW01600 standard; protein; 498 AA.  
 XX AAW01600;  
 AC  
 DT 17-MAR-1997 (first entry)  
 XX  
 DE Thrombomodulin TME456 protein.  
 XX  
 KW thrombomodulin; nervous interference; treatment; medulla injury.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP08283174-A.  
 XX  
 PD 29-OCT-1996.  
 XX  
 PF 11-APR-1995; 95JP-00085202.  
 XX  
 PR 11-APR-1995; 95JP-00085202.  
 XX  
 PA (ASAH ) ASAH KASEI KOGYO KK.  
 XX  
 DR WPI; 1997-017314/02.  
 XX  
 PT Treating agent for nervous interference after medulla injury - contains  
 PT thrombomodulin as active component.  
 XX  
 PS Claim 2; Page 5-6; 6pp; Japanese.  
 XX  
 CC AAW01599-600 are fragments of human thrombomodulin. They can be used in  
 CC treatment of nervous interference after medulla injury. In an example, a  
 CC model of medulla injury was prepd. by compressing a male Wistar rat of  
 CC 200-250 g body wt., with a 20g weight for 20 mins. TMD123 dissolved in  
 CC physiological saline water was injected into the tail vein 30 mins.  
 CC before injury. Movement function was evaluated 24 hrs. after injury. MPO  
 CC activity in medulla tissue was measured 3 hrs. after injury as an index

CC of neutrophile accumulation. The Tarlov score was 1.50 compared to 0.77  
 CC for the control, and complete paralysis was 15.4 percent compared to 35.3  
 CC percent for the control  
 XX  
 SQ Sequence 498 AA;

Query Match 100.0%; Score 2830; DB 2; Length 498;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 DB 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 QY 61 GGVGRRRLWIGLQPLPGCGDKRLGRLGPGFOWTGDNNTSYSRWARLDLNGAPLCGLPLCV 120  
 DB 61 GGVGRRRLWIGLQPLPGCGDKRLGRLGPGFOWTGDNNTSYSRWARLDLNGAPLCGLPLCV 120  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGMAAAAVSITYGTPFA 180  
 DB 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGMAAAAVSITYGTPFA 180  
 QY 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 DB 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMCTGYRLAA 300  
 DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMCTGYRLAA 300  
 QY 301 DOHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
 DB 301 DOHRCEDVDDCILEPSPCQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
 QY 361 QPLNQTSLYLCVCAEGEAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPBGYILDDGFI 420  
 DB 361 QPLNQTSLYLCVCAEGEAPIPHEPHRCQMFNCOTACPADCDPNTQASCEPBGYILDDGFI 420  
 QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGSEPPPS 480  
 DB 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGSEPPPS 480  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 481 PTPGSTLTTPPAVGLVHSG 498

RESULT 3  
 AAR22016  
 ID AAR22016 standard; protein; 516 AA.  
 XX  
 AC AAR22016;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 03-JUL-1992 (first entry)  
 XX  
 DE Truncated human thrombomodulin encoded by plasmid psv2TMD1.  
 XX  
 KW Thrombin binding site; blood clotting; TMD1 deleter.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP474273-A.  
 XX  
 PD 11-MAR-1992.  
 XX  
 PF 05-AUG-1991; 91EP-00202009.  
 XX  
 PR 03-AUG-1990; 90TP-00204978.  
 XX  
 PA (ASAH ) ASAH KASEI KOGYO KK.  
 XX  
 PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;

XX DR WPI; 1992-081820/11.  
 XX PT New polypeptide inhibits blood coagulation and platelet aggregation -  
 XX PT promotes thrombin catalysed protein C activation for treating myocardial  
 XX PT infarction, thrombosis, embolism, etc.  
 XX PS Example 1; Page 17; 112pp; English.  
 XX CC Plasmid M13mpl19TMJ3 (see AAR22013) was used as template for site-directed  
 CC mutagenesis. A 177bp fragment was deleted using the "TMDI deleter"  
 CC oligonucleotide to give plasmid M13TMD1 which encodes the first 516 N-  
 CC terminal amino acids of human thrombomodulin. Plasmid M13TMD1 was  
 CC completely digested with HindIII and BamHI and a TMDI fragment of ca.  
 CC 1700bp was isolated. The fragment was ligated to HindIII- and BglII-cut  
 CC plasmid pSV2-dhfr to give the recombinant plasmid pSV2TMD1. See AAR22014-  
 CC R22022 and AAR25072. (Updated on 25-MAR-2003 to correct PA field.)  
 XX SQ Sequence 516 AA;

Query Match 100.0%; Score 2830; DB 2; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQPPGCGDKPLRGFOWVTGDNNTSYSRWRLDNGAPLCGLCV 120  
 Db 79 GGVGRRRLWIGLQPPGCGDKPLRGFOWVTGDNNTSYSRWRLDNGAPLCGLCV 138  
 QY 121 AVSAEATVPSEPIWEEQCEVKADGFLCEFPFATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 139 AVSAEATVPSEPIWEEQCEVKADGFLCEFPFATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 240  
 Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 258  
 QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCHEFCVNPDPQSGYSCMCEGYRLAA 300  
 Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCHEFCVNPDPQSGYSCMCEGYRLAA 318  
 QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHYCNVYDLVDGECVPEVDFCFRANCEYQC 360  
 Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFCHYCNVYDLVDGECVPEVDFCFRANCEYQC 378  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCPNTQASCECEGYLLDDGFI 420  
 Db 379 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCPNTQASCECEGYLLDDGFI 438  
 QY 421 CTDIDCEGCGFSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSDSGSGEPSPS 480  
 Db 439 CTDIDCEGCGFSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSDSGSGEPSPS 498  
 QY 481 PTPGSTLTTPPAVLVHSG 498  
 Db 499 PTPGSTLTTPPAVLVHSG 516

RESULT 4  
 AAY09347  
 ID AAY09347 standard; protein; 516 AA.  
 XX AC  
 XX AC AAY09347;  
 XX DT  
 XX DT 08-JUL-1999 (first entry)  
 XX DE Human thrombomodulin SEQ ID NO:1.  
 XX KW Human; thrombomodulin; aqueous parenteral solution; storage;  
 KW distribution; acute coronary syndrome; thrombosis; embolism; diabetes.

XX OS Homo sapiens.  
 XX PN WO9918994-A1.  
 XX XX 22-APR-1999.  
 XX PF 13-OCT-1998; 98WO-JP004609.  
 XX PR 15-OCT-1997; 97JP-00281659.  
 XX PR 11-NOV-1997; 97JP-00308523.  
 XX PA (ASAH) ASAH KASEI KOGYO KK.  
 XX PI Yui M, Yokozawa A, Murata T, Tsuruta K, Shimizu H;  
 XX WPI; 1999-277444/23.  
 XX N-ESDB; AAX55879.

Stable aqueous parenteral thrombomodulin solution - comprising buffer and  
 surfactant, useful for treating acute coronary syndrome, thrombosis,  
 embolism, and diabetes.

Claim 6; Page 87-89; 97pp; Japanese.

The present invention describes a method for maintaining the quality of  
 an aqueous parenteral solution of thrombomodulin comprising buffer and  
 surfactant aseptically filled in a case or syringe. Maintaining the  
 quality of an aqueous, parenteral thrombomodulin solution is  
 characterised in that the solution: (a) comprises soluble thrombomodulin,  
 a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
 filled into: (i) a case or (ii) a syringe without any empty space; and  
 (c) is kept in liquid form in storage and distribution and not frozen or  
 freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
 (such as myocardial infarction, unstable angina and coronary artery  
 blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
 vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
 Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic  
 lupus erythematosus or Barrett's syndrome), multiple organ failure,  
 disseminated intravascular coagulation, transient ischaemia, diabetes,  
 liver veno-occlusive diseases and deep vein thrombosis. The composition  
 is stable for a long period of time and can be stored and distributed in  
 ready to use form avoiding the problems of dissolution and accuracy when  
 preparing on demand. The present sequence represents human thrombomodulin

SQ Sequence 516 AA;

Query Match 100.0%; Score 2830; DB 2; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQPPGCGDKPLRGFOWVTGDNNTSYSRWRLDNGAPLCGLCV 120  
 Db 79 GGVGRRRLWIGLQPPGCGDKPLRGFOWVTGDNNTSYSRWRLDNGAPLCGLCV 138  
 QY 121 AVSAEATVPSEPIWEEQCEVKADGFLCEFPFATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 139 AVSAEATVPSEPIWEEQCEVKADGFLCEFPFATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 240  
 Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 258  
 QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCHEFCVNPDPQSGYSCMCEGYRLAA 300  
 Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCHEFCVNPDPQSGYSCMCEGYRLAA 318  
 QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHYCNVYDLVDGECVPEVDFCFRANCEYQC 360



Db 319 DQHRCEVDVDCILBSPQPCQRCVNTQGGFECHCYENYDLVDGECVEPVDPDFRANCEYQC 378  
 QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCPCPGYILDDGFI 420  
 Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCPCPGYILDDGFI 438  
 QY 421 CTIDIDECENGFCGCVCHNLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGSGGEPPEPS 480  
 Db 439 CTIDIDECENGFCGCVCHNLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGSGGEPPEPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 5  
 AAY83934  
 ID AAY83934 standard; protein; 516 AA.  
 XX AC AAY83934;  
 DT 28-JUL-2000 (first entry)  
 DE Human thrombomodulin TMD protein.  
 XX Human; thrombomodulin; vasculitis; protein C; thrombin.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN JP2000053582-A.  
 XX 22-FEB-2000.  
 XX 06-AUG-1998; 98JP-00222688.  
 XX 06-AUG-1998; 98JP-00222688.  
 XX (ASAH) ASAHI KASEI KOGYO KK.  
 DR WPI; 2000-353249/31.  
 DR N-PSDB; AAA10027.  
 XX Treating agent for vasculitis contains peptide which promotes activation  
 PT of proterin C by thrombin.  
 XX Claim 4; Page 10-12; 18pp; Japanese.  
 XX This sequence represents a human thrombomodulin protein. The invention  
 CC relates to a treating agent for vasculitis containing a peptide which  
 CC promotes activation of protein C by thrombin  
 XX Sequence 516 AA;  
 SQ

Query Match 100.0%; Score 2830; DB 3; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149; Indels 0; Gaps 0;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 78  
 QY 61 GGVRRRLWGLQLPPCGGPKLGLPGRGQWVTGDNNTSYRWARLDLNGAPLQPLCV 120  
 Db 79 GGVRRRLWGLQLPPCGGPKLGLPGRGQWVTGDNNTSYRWARLDLNGAPLQPLCV 138  
 QY 121 AVSAAEATVPSEPIWEQQCEVADGFLCFRHPFATCRPLAVEPFGAAAVSITYGTPEA 180  
 Db 139 AVSAAEATVPSEPIWEQQCEVADGFLCFRHPFATCRPLAVEPFGAAAVSITYGTPEA 198  
 QY 181 ARGADFQALPVGSSAAVAPLGLQLMCTAPPFGAVQGHWAREAPGAWDCSVENGCGEACNA 240  
 Db 199 ARGADFQALPVGSSAAVAPLGLQLMCTAPPFGAVQGHWAREAPGAWDCSVENGCGEACNA 258

QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318  
 QY 301 DOHRCEVDVDCILBSPQPCQRCVNTQGGFECHCYENYDLVDGECVEPVDPDFRANCEYQC 360  
 Db 319 DOHRCEVDVDCILBSPQPCQRCVNTQGGFECHCYENYDLVDGECVEPVDPDFRANCEYQC 378  
 QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCPCPGYILDDGFI 420  
 Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCPCPGYILDDGFI 438  
 QY 421 CTIDIDECENGFCGCVCHNLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGSGGEPPEPS 480  
 Db 439 CTIDIDECENGFCGCVCHNLPGTFECICGPDLSALVRHIGTDCDCKVDDGSGSGGEPPEPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 6  
 AAY69529  
 ID AAY69529 standard; protein; 516 AA.  
 XX AC AAY69529;  
 DT 10-APR-2000 (first entry)  
 DE Human thrombomodulin variant, SEQ ID NO:3.  
 XX Thrombomodulin; TM; soluble; affinity purification; cation exchange;  
 KW anticoagulant; thrombolytic.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN JP113411990-A.  
 XX 14-DEC-1999.  
 XX 23-MAR-1999; 99JP-00077518.  
 XX 30-MAR-1998; 98JP-00084389.  
 XX (ASAH) ASAHI KASEI KOGYO KK.  
 DR WPI; 2000-101696/09.  
 DR N-PSDB; AAZ55965.  
 XX Preparation of highly pure soluble thrombomodulin - used as an antillood  
 PT coagulation agent and a thrombolytic agent.  
 XX Claim 9; Page 30-32; 36pp; Japanese.  
 XX The invention relates to a novel method for the preparation of highly  
 CC pure soluble thrombomodulin (TM) containing substantially no serum-  
 CC derived or antibody-derived substance. The method comprises isolating  
 CC soluble TM from unpurified serum via affinity purification using an anti-  
 CC TM antibody. The soluble TM is then treated with a cation exchanger  
 CC at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to 4, and the  
 CC fraction containing the soluble TM is isolated. The soluble TM can be  
 CC used as a blood anticoagulation agent and a thrombolytic agent. This  
 CC sequence represents a human thrombomodulin variant, designated SEQ ID  
 CC NO:3  
 XX Sequence 516 AA;  
 SQ

Query Match 100.0%; Score 2830; DB 3; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-149; Indels 0; Gaps 0;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 78  
 QY 61 GGVRRRLWGLQLPPCGGPKLGLPGRGQWVTGDNNTSYRWARLDLNGAPLQPLCV 120  
 Db 79 GGVRRRLWGLQLPPCGGPKLGLPGRGQWVTGDNNTSYRWARLDLNGAPLQPLCV 138  
 QY 121 AVSAAEATVPSEPIWEQQCEVADGFLCFRHPFATCRPLAVEPFGAAAVSITYGTPEA 180  
 Db 139 AVSAAEATVPSEPIWEQQCEVADGFLCFRHPFATCRPLAVEPFGAAAVSITYGTPEA 198  
 QY 181 ARGADFQALPVGSSAAVAPLGLQLMCTAPPFGAVQGHWAREAPGAWDCSVENGCGEACNA 240  
 Db 199 ARGADFQALPVGSSAAVAPLGLQLMCTAPPFGAVQGHWAREAPGAWDCSVENGCGEACNA 258

|     |  |     |
|-----|--|-----|
| 19  | APAEPGGGQCVEHDCFALYEGPATFNLNASQICDGLRGHLMTVRSVAADVISLLNGD    | 78  |
| 61  | GGVRRRLMIGLQLPFGCGDPKRLGFLRGFQWVTGDMNNTYSRWARLDLNGAPLCGPLCV  | 120 |
| 79  | GGVRRRLMIGLQLPFGCGDPKRLGFLRGFQWVTGDMNNTYSRWARLDLNGAPLCGPLCV  | 138 |
| 121 | AVSAAEATVPSPIWEEQCEVKADGFTCEHFHPATCRPLAVEPGAAAAVSYTGTTPFA    | 180 |
| 139 | AVSAAEATVPSPIWEEQCEVKADGFTCEHFHPATCRPLAVEPGAAAAVSYTGTTPFA    | 198 |
| 181 | ARGADFOALPVGSSAAVPLGILQMLCTAPPAGAVQGHWAREAPGAWDCSVNGGCEHACNA | 240 |
| 199 | ARGADFOALPVGSSAAVPLGILQMLCTAPPAGAVQGHWAREAPGAWDCSVNGGCEHACNA | 258 |
| 241 | IFGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNPDPDGGSYSCMCTGTYRLAA | 300 |
| 259 | IFGAPRCQCPAGAALQADGRSCTASATQSCNDLCEHFCVNPDPDGGSYSCMCTGTYRLAA | 318 |
| 301 | DQHRCEDVDVDCILLESPCQRCVNTQGGFGECHCVNTDLVDGCEVPDPCFRANCEYQC   | 360 |
| 319 | DQHRCEDVDVDCILLESPCQRCVNTQGGFGECHCVNTDLVDGCEVPDPCFRANCEYQC   | 378 |
| 361 | QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI | 420 |
| 379 | QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI | 438 |
| 421 | CTDIDBCENGFGCSGVCHNLPTFECICGPDASALVRHIGTDCDSKVDGDSGSGPPPS    | 480 |
| 439 | CTDIDBCENGFGCSGVCHNLPTFECICGPDASALVRHIGTDCDSKVDGDSGSGPPPS    | 498 |
| 481 | PTPGSTLTTPPAVGLVHSG  | 498 |
| 499 | PTPGSTLTTPPAVGLVHSG  | 516 |

|          |   |
|----------|---|
| RESULT 7 |   |
| ABU08701 |   |
| ID       | ABU08701 standard; protein; 516 AA.                                       |
| XX       |   |
| XX       |   |
| AC       | ABU08701;   |
| XX       |   |
| XX       |   |
| DT       | 17-JUN-2003 (first entry)   |
| XX       |   |
| DE       | Thrombomodulin-containing composition associated protein #1.              |
| XX       |   |
| XX       | Thrombomodulin; antiplatelet; anticoagulant; thrombolytic;                |
| KW       | blood coagulation-fibrinolysis system; vascular obstruction;              |
| KW       | myocardial infarction; acute coronary syndrome; brain infarction; sepsis; |
| KW       | cardiant; cerebroprotective; antibacterial; immunosuppressive;            |
| KW       | anticoagulant; thrombolytic.  |
| XX       |   |
| OS       | Unidentified.   |
| XX       |   |
| XX       | WO2003013606-A1.  |
| PN       |   |
| XX       |   |
| XX       | 20-FEB-2003.  |
| PD       |   |
| XX       |   |
| XX       |   |
| PF       | 07-AUG-2002; 2002WO-JP008058.   |
| XX       |   |
| XX       |   |
| PR       | 08-AUG-2001; 2001JP-00241371.   |
| XX       |   |
| XX       | (ASAH ) ASAHI KASEI KOGYO KK.   |
| PA       |   |
| XX       |   |
| PI       | Suzuki H, Mohri M;  |
| XX       |   |
| XX       | WPI; 2003-248241/24.  |
| DR       | N-PSDB; ABX94087.   |
| XX       |   |
| XX       |   |
| PT       | Thrombomodulin-containing medicinal compositions useful for treating or   |
| PT       | preventing e.g. myocardial infarction, brain infarction, acute coronary   |
| PT       | syndrome and sepsis.  |
| XX       |   |
| PS       | Claim 18; Page 29-31; 48pp; Japanese.                                     |

|    |   |
|----|---|
|    | The invention describes drug compositions containing thrombomodulin, and<br>an antiplatelet, an anticoagulant or/ and a thrombolytic other than<br>thrombomodulin. The drug compositions are for preventing or treating<br>diseases or syndromes due to abnormality in the blood coagulation-<br>fibrinolysis system or/and vascular obstruction, including myocardial<br>infarction, acute coronary syndrome, brain infarction and sepsis,<br>particularly as combination drug compositions. The compositions have high<br>therapeutic efficacy. This is the amino acid sequence of a thrombomodulin-<br>containing composition associated protein |
| XX | SQ Sequence 516 AA;   |
|    | Query Match 100.0%; Score 2830; DB 6; Length 516;<br>Best Local Similarity 100.0%; Pred. No. 1.1e-149;<br>Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0   |
| QY | 1 APAEPQGGSCVCEHDCFALYPGPATFTFNASQICDGLRGHLMTVRSSVAADVISILLNGD 60   |
| Db | 19 APAEPQGGSCVCEHDCFALYPGPATFTFNASQICDGLRGHLMTVRSSVAADVISILLNGD 78  |
| QY | 61 GGVGERRRLWIGLQLPPGCGDPKRLGPLRGFOWTGDNNNTSYSRWARLDINGAPLCGPLCV 120  |
| Db | 79 GGVGERRRLWIGLQLPPGCGDPKRLGPLRGFOWTGDNNNTSYSRWARLDINGAPLCGPLCV 138  |
| QY | 121 AVSAAEATVPSEPTWEEOOCEVKADGFLICEHPFPATCHPLAVEPGAAAANSITTGTTPA 180  |
| Db | 139 AVSAAEATVPSEPTWEEOOCEVKADGFLICEHPFPATCHPLAVEPGAAAANSITTGTTPA 198  |
| QY | 181 ARGADFQALPVGSAAVAPLGILQMCTAPPAGVQHWAREAPGAWDCSVENGCCBHACNA 240  |
| Db | 199 ARGADFQALPVGSAAVAPLGILQMCTAPPAGVQHWAREAPGAWDCSVENGCCBHACNA 258  |
| QY | 241 IPGAPRCQPAGAALQADGRSCTASATOSCNDLCEHFVCVNPDPQSGYSMCETGYRLAA 300  |
| Db | 259 IPGAPRCQPAGAALQADGRSCTASATOSCNDLCEHFVCVNPDPQSGYSMCETGYRLAA 318  |
| QY | 301 DQHCEDVDVCILLEPSPCPQCYNVTQGGEFCCHCYPNYDLVDGECVEPVDPCFRANCEYCQ 360   |
| Db | 319 DQHCEDVDVCILLEPSPCPQCYNVTQGGEFCCHCYPNYDLVDGECVEPVDPCFRANCEYCQ 378   |
| QY | 361 QPLNQTSLYLCVCSGGFAPIPEPHRCQMFNCQTACPADCENTQASCCEPGVILLDDGFI 420   |
| Db | 379 QPLNQTSLYLCVCSGGFAPIPEPHRCQMFNCQTACPADCENTQASCCEPGVILLDDGFI 438   |
| QY | 421 CTDIDECENGGFGCSGVCHNLPGTFTECIPGPSALSALVRHITGDTCDSGKVDPGDSGSGEPPPS 480   |
| Db | 439 CTDIDECENGGFGCSGVCHNLPGTFTECIPGPSALSALVRHITGDTCDSGKVDPGDSGSGEPPPS 498   |
| QY | 481 PTPGSTLTTPPAVLVHSG 498<br>  |
| Db | 499 PTPGSTLTTPPAVLVHSG 516<br>  |

|          |  |
|----------|--|
| RESULT 8 |  |
| ADA37271 |  |
| ID       | ADA37271 standard; protein; 516 AA.                            |
| XX       |  |
| AC       | ADA37271;  |
| XX       |  |
| DT       | 20-NOV-2003 (first entry)                                      |
| XX       |  |
| DE       | Human soluble thrombomodulin protein SEQ ID NO:1.              |
| XX       |  |
| KW       | high-concentration preparation; soluble thrombomodulin; human; |
| KW       | thrombomodulin.  |
| XX       |  |
| OS       | Homo sapiens.  |
| XX       |  |
| FN       | WO2003061687-A1.   |
| XX       |  |
| PD       | 31-JUL-2003.   |
| XX       |  |

```

PF 17-JAN-2003; 2003WO-JP000339.
XX
PR 18-JAN-2002; 2002JP-00009951.
XX
PA (ASAH ) ASahi KASEI KK.
XX
PI Nishio F;
XX
DR WPI; 2003-712487/67.
XX
DR N-PSDB; ADA37272.
XX
PT High concentration stable low-foaming soluble thrombomodulin preparation
PT for pharmaceutical use.
XX
XX Claim 18; Page 96-98; 113pp; Japanese.
PS
CC The present invention describes the high-concentration preparation of
CC soluble thrombomodulin. The present sequence represents a human soluble
CC thrombomodulin, which is given in the exemplification of the present
CC invention.
XX
SQ Sequence 516 AA;

Query Match          100.0%; Score 2830; DB 7; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.1e-149;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78
QY 61 GGVGRRRLWIGLQPLPGCGDPRKRLGPRFQWVTGDNNTSYSRWARLDNGAPLPGPLCV 120
Db 79 GGVGRRRLWIGLQPLPGCGDPRKRLGPRFQWVTGDNNTSYSRWARLDNGAPLPGPLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFHPCFATCRPLAVEPGAAAAVSIYGTPTFA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFHPCFATCRPLAVEPGAAAAVSIYGTPTFA 198
QY 181 ARGADFOALFVGSAAVAPLGLQIMCTAPPGVQGHWAREAPGAWDCSVENGCGCEHACNA 240
Db 199 ARGADFOALFVGSAAVAPLGLQIMCTAPPGVQGHWAREAPGAWDCSVENGCGCEHACNA 258
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318
QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYENYDLVDGECVEPVDPFRANCEYQC 360
Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYENYDLVDGECVEPVDPFRANCEYQC 378
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCMFCNQACFADCDPNTQASCCEPCGYILDDGFI 420
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCMFCNQACFADCDPNTQASCCEPCGYILDDGFI 438
QY 421 CTDIDECENGFCSCVCHNLPGTFECICGPDLSALVRHIGTDCDQSKVDGSGSGSEPPPS 480
Db 439 CTDIDECENGFCSCVCHNLPGTFECICGPDLSALVRHIGTDCDQSKVDGSGSGSEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 9
AAR11534
ID AAR11534 standard; protein; 575 AA.
XX
XX AAR11534;
AC
XX
XX
DT 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 17-JUN-1991 (first entry)

```

```

XX Human thrombomodulin type II polypeptide.
DE
XX Thrombomodulin; chondroitin; protein C; thrombin; thrombolysis;
KW chondroitin sulphate; anticoagulant; myocardial infarction.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT Peptide /label= signal peptide
FT Peptide 19..365
FT Peptide /label= opt. N-terminal sequence
FT Peptide 366..480
FT Peptide /label= human thrombomodulin active fragment
FT Peptide 481..516
FT Peptide /label= opt. C-terminal sequence
XX
PN W09104276-A.
XX
PD 04-APR-1991.
XX
PF 25-SEP-1989; 89JP-00246270.
XX
PR 25-SEP-1989; 89JP-00246270.
XX
PA (ASAH ) ASahi KASEI KOGYO KK.
XX
PI Yamamoto S, Gomi K, Ogawa K;
XX WPI; 1991-117478/16.
XX
DR New human thrombomodulin polypeptide contg. chondroitin - has
PT anticoagulant platelet agglutination and thrombolytic activity.
XX
PS Disclosure; Fig 1; 80pp; Japanese.
XX
CC This human thrombomodulin polypeptide comprises a 115 residue active
CC fragment with opt. flanking sequences (N-terminal= 347 amino acids; C-
CC terminal= 36 amino acids). Additionally it has an N-terminal signal
CC sequence and an attached sugar chain, contg. chondroitin or chondroitin
CC sulphate. It promotes the activation of protein C by thrombin and has
CC good anticoagulant, platelet aggregation inhibition and thrombolytic
CC activities. It can therefore be used for treating cardiovascular
CC diseases, eg arteriosclerosis or myocardial infarction. (Updated on 09-
CC JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
SQ Sequence 575 AA;

Query Match          100.0%; Score 2830; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.2e-149;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78
QY 61 GGVGRRRLWIGLQPLPGCGDPRKRLGPRFQWVTGDNNTSYSRWARLDNGAPLPGPLCV 120
Db 79 GGVGRRRLWIGLQPLPGCGDPRKRLGPRFQWVTGDNNTSYSRWARLDNGAPLPGPLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFHPCFATCRPLAVEPGAAAAVSIYGTPTFA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFHPCFATCRPLAVEPGAAAAVSIYGTPTFA 198
QY 181 ARGADFOALFVGSAAVAPLGLQIMCTAPPGVQGHWAREAPGAWDCSVENGCGCEHACNA 240
Db 199 ARGADFOALFVGSAAVAPLGLQIMCTAPPGVQGHWAREAPGAWDCSVENGCGCEHACNA 258
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318

```

QY 301 DOHRCEDVDDCILEPSPQPCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQC 360  
 DB 319 DOHRCEDVDDCILEPSPQPCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQC 378  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCPDNTQASCECPGYILDDGFI 420  
 DB 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCPDNTQASCECPGYILDDGFI 438  
 QY 421 CTIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGGSEPPPS 480  
 DB 439 CTIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGGSEPPPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 499 PTPGSTLTTPPAVGLVHSG 516  
  
 RESULT 10  
 AAR41806  
 ID AAR41806 standard; peptide; 575 AA.  
 XX AC AAR41806;  
 XX 30-MAR-1994 (first entry)  
 DT 30-MAR-1994 (first entry)  
 DB Human thrombomodulin.  
 XX Transformation; fungus; blood coagulation; prevention; platelet;  
 KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;  
 KW site-directed mutagenesis; promotion; protein C; activation; thrombin.  
 XX  
 OS Homo sapiens.  
 XX JP05213998-A.  
 XX 24-AUG-1993.  
 XX 02-AUG-1991; 91JP-00282369.  
 XX 03-AUG-1990; 90JP-00204978.  
 XX 30-JUL-1991; 91JP-00189984.  
 XX (ASAHI ) ASAHI CHEM IND CO LTD.  
 XX WPI; 1993-299652/38.  
 XX Novel polypeptide obtd. by culturing transformed fungus - having blood  
 PT coagulation preventing, platelet aggregation preventing and thrombolytic  
 PT activities.  
 XX Disclosure; Fig 55; 65pp; Japanese.  
 XX  
 CC Novel polypeptides, obtd. by culturing transformed fungus, have blood  
 CC coagulation preventing, platelet aggregation preventing and thrombolytic  
 CC activities. In an example, plasmid M13mpl9TMD3 (constructed from pSV2TMJ2  
 CC (ATCC 67238) contg. a human thrombomodulin sequence) was subjected to  
 CC site directed mutagenesis to prepare pSV2TMD7. Plasmid pSV2TMD7 was  
 CC transfected to COS-1 cells. The activity of promoting protein C  
 CC activation by thrombin of the peptide produced by the transformed COS-1  
 CC cell was measured. The amt. of the peptide was determined  
 XX  
 SQ Sequence 575 AA;  
 Query Match 100.0%; Score 2830; DB 2; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-149; Indels 0; Gaps 0;  
 Matches 498; Conservative 0; Mismatches 0;  
 QY 1 APAEPQGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRRSSVAADVISLLNGD 60  
 DB 19 APAEPQGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRRSSVAADVISLLNGD 78  
 QY 61 GGVGRRRLWIGLQLPPGCGDPKRLGRLGFWQVTDGNTSYSRWARLDLNGAPLCGPLCV 120

DB 79 GGVGRRRLWIGLQLPPGCGDPKRLGRLGFWQVTDGNTSYSRWARLDLNGAPLCGPLCV 138  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEPHFPATCRPLAVEPGAAAAAASVITYGTTPA 180  
 DB 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEPHFPATCRPLAVEPGAAAAAASVITYGTTPA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHCNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHCNA 258  
 QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCETGYRLAA 300  
 DB 259 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCETGYRLAA 318  
 QY 301 DOHRCEDVDDCILEPSPQPCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQC 360  
 DB 319 DOHRCEDVDDCILEPSPQPCVNTQGGFECHCYPNYDLVDGECVEPVDPCFRANCEYQC 378  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCPDNTQASCECPGYILDDGFI 420  
 DB 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCPDNTQASCECPGYILDDGFI 438  
 QY 421 CTIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGGSEPPPS 480  
 DB 439 CTIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGGSEPPPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 499 PTPGSTLTTPPAVGLVHSG 516  
  
 RESULT 11  
 AAR43031  
 ID AAR43031 standard; protein; 575 AA.  
 XX AC AAR43031;  
 XX 25-MAR-2003 (revised)  
 DT 16-MAY-1994 (first entry)  
 XX Human thrombomodulin.  
 XX Anticoagulant; platelet aggregation inhibitor; protein C; activation;  
 KW thrombin; thrombomodulin; coagulation disorder; thrombosis;  
 KW myocardial infarction; embolism; telangiectaphraxis;  
 KW arteriosclerosis obliterans; disseminated intravascular coagulation; DIC;  
 KW angina pectoris; gestosis; transient ischaemic attack.  
 XX Homo sapiens.  
 XX WO9322447-A1.  
 XX 11-NOV-1993.  
 XX 30-APR-1993; 93WO-JP000578.  
 XX 01-MAY-1992; 92JP-00112903.  
 XX (ASAHI ) ASAHI KASEI KOGYO KK.  
 XX Zushi M, Kondo S, Toma K;  
 XX WPI; 1993-368806/46.  
 XX Peptide with anticoagulant and platelet aggregation inhibitor activity -  
 PT which promotes protein C activation by thrombin and is useful in treating  
 PT coagulation disorders e.g. thrombosis.  
 XX Disclosure; Fig 1; 84pp; Japanese.  
 CC New peptides (see AAR50069) are inhibitors of the blood coagulation and  
 CC platelet aggregation activities of thrombin and promote the protein-C

activation effect of thrombin. They can be produced efficiently in pure form by culture of appropriate transformants, and are useful in treatment of circulatory disorders such as myocardial infarction, thrombosis, embolism, telangiectasis, arteriosclerosis obliterans, disseminated intravascular coagulation, angina pectoris, gestosis and transient ischaemic attack. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 575 AA;

Query Match 100.0%; Score 2830; DB 2; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-149;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 60  
 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 78  
 61 GGVGRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCPLCV 120  
 79 GGVGRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCPLCV 138  
 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAVSITGTPFA 180  
 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAVSITGTPFA 198  
 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCCCHACNA 240  
 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCCCHACNA 258  
 241 IPGAPRCOPAGALQADGRSCTASATOSCNLDLCEHFCVNPDPQGSYSCMCTGYRLAA 300  
 259 IPGAPRCOPAGALQADGRSCTASATOSCNLDLCEHFCVNPDPQGSYSCMCTGYRLAA 318  
 301 DQHRCEVDVDCILESPQPCQVNTQGGFECHCYNYDLVDGECVPEVDPGFRANCEYQC 360  
 319 DQHRCEVDVDCILESPQPCQVNTQGGFECHCYNYDLVDGECVPEVDPGFRANCEYQC 378  
 361 QPLNQTSLVCVAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYIIDDGFI 420  
 379 QPLNQTSLVCVAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYIIDDGFI 438  
 421 CTDIDECENGFCSCVCHNLGTFECICGPPSALVRHGTGDCSGKVVDGSGSGGEPSPS 480  
 439 CTDIDECENGFCSCVCHNLGTFECICGPPSALVRHGTGDCSGKVVDGSGSGGEPSPS 498  
 481 PTPGSTLTTPPAVGLVHSG 498  
 499 PTPGSTLTTPPAVGLVHSG 516

# RESULT 12

AAE17521  
 ID AAE17521 standard; protein; 575 AA.

AC AAE17521;

XX AAE17521;

XX 22-APR-2002 (first entry)

XX Human full-length native thrombomodulin (TM).

Human; thrombomodulin; TM analogue; protein C; thrombotic disease; antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor; TAFI; myocardial infarction; disseminated intravascular coagulation; DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina; acute respiratory distress syndrome; cancer; toxemia; septicemia; circulatory disorder; coronary embolism; pulmonary embolism; systemic coagulation disorder; immunosuppressive; cerebroprotective; pregnancy; anticoagulant; thrombolytic; venous occlusive condition; antibacterial; cardiant.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..18  
 FT Protein /label= Signal\_peptide 19..575  
 FT Domain /note= "Human mature thrombomodulin" 19..244  
 FT Domain /label= N-terminal\_domain 245..480  
 FT Modified-site /note= "6 EGF-like domains" 481..515  
 FT Domain /note= "O-linked glycosylation site" 516..539  
 FT Domain /label= Transmembrane\_domain 540..575  
 FT Domain /label= Cytoplasmic\_domain

PN WO200198352-A2.

PD 27-DEC-2001.

PF 20-JUN-2001; 2001WO-US019590.

PR 21-JUN-2000; 2000US-0213678P.

PR 12-JUN-2001; 2001US-00880484.

XX (SCHD ) SCHERING AG.

XX Light D, Morser MJ, Nagashima M;

XX WPI; 2002-154622/20.

XX N-PSDB; AAD28182.

Thrombomodulin analog with greater ability to potentiate thrombin-mediated activation of protein C and lesser ability to potentiate thrombin-mediated activation of thrombin-activatable fibrinolysis inhibitor for treating thrombotic disease.

Claim 1; Fig 4; 53pp; English.

The invention relates to the design, production and use of thrombomodulin (TM) analogues that have the ability to enhance the thrombin-mediated activation of protein C which have a significantly reduced ability to promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI). TM analogue is useful for treating thrombotic disease in human, by administering it or a polynucleotide encoding it to the patient and expressing the analogue in vivo. Composition comprising TM analogue is useful for treating a thrombotic disease or condition in mammals. TM analogue is useful in antithrombotic therapy. TM analogue is also useful for treating diseases in which thrombus formation plays a significant etiological role including myocardial infarction, disseminated intravascular coagulation (DIC), deep vein thrombosis, septic shock, pulmonary embolism, acute respiratory distress syndrome, unstable angina and other arterial or venous occlusive conditions. TM analogue is also useful for treating various circulatory disorders including coronary or pulmonary embolism, stroke and systemic coagulation disorders including DIC which is often associated with septicemia, certain cancers and toxemia of pregnancy. The present sequence is human full-length native thrombomodulin

SQ Sequence 575 AA;

Query Match 100.0%; Score 2830; DB 5; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-149;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSVAADVLSLLNGD 78  
 QY 61 GGVGRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCPLCV 120  
 Db 79 GGVGRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWRLDLNGAPLCPLCV 138  
 QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAVSITGTPFA 180

Db 139 AVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSIYGTTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 258  
QY 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDQPGSYSCMCETGYRLAA 300  
Db 259 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDQPGSYSCMCETGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECYCNVNDLVGECVEPVPDPCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECYCNVNDLVGECVEPVPDPCFRANCEYQC 378  
QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQFCNOTACPADCDPNTQASCEPGEYIILDDGFI 420  
Db 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQFCNOTACPADCDPNTQASCEPGEYIILDDGFI 438  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSDSGSGEPPEPS 480  
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSDSGSGEPPEPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 13

AAE23026  
ID AAE23026 standard; protein; 575 AA.

XX AC AAE23026;

XX DT 21-AUG-2002 (first entry)

XX DE Human thrombomodulin protein.

XX KW Human; thrombomodulin; analogue; neurologic trauma; vulnery;  
KW spinal cord injury; neurological damage; therapy; tranquilliser;  
KW neuroprotective.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Domain 1..226 /note= "N-terminal domain"  
FT Peptide 1..18 /label= Signal\_peptide  
FT Protein 19..575 /note= "Human mature thrombomodulin"  
FT Domain 227..462 /note= "6 EGF-like domain"  
FT Modified-site 463..497 /note= "O-linked glycosylation site"  
FT Domain 498..521 /note= "Transmembrane domain"  
FT Domain 522..557 /note= "Cytoplasmic domain"

XX PN WO200217953-A2.

XX PD 07-MAR-2002.

XX PF 31-AUG-2001; 2001WO-US041930.

XX PR 31-AUG-2000; 2000US-0229714P.

XX PR 23-AUG-2001; 2001US-00938405.

XX PA (SCHD ) SCHERING AG.

XX PI Festoff BW, Morser MJ;

XX XX

DR WPI: 2002-415551/44.  
XX N-PSDB; AAD36901.  
PT Treating neurological damage resulting from spinal cord injury in humans  
PT by administering a soluble, recombinant thrombomodulin analog which is  
PT resistant to oxidation.  
XX  
PS Claim 1; Fig 1; 31pp; English.  
XX  
CC The invention relates to a method of using thrombomodulin analogues in  
CC the treatment of the neurologic trauma associated with spinal cord injury  
CC in mammals. The method involves administering an oxidation resistant,  
CC soluble, recombinant thrombomodulin analogue, where the methionine at  
CC position 388 has been replaced with a leucine, and the analogue is  
CC numbered in accordance with native thrombomodulin. The method and  
CC thrombomodulin analogue are useful for treating neurological damage  
CC resulting from spinal cord injury in human. The present sequence is human  
CC thrombomodulin protein  
XX  
SQ Sequence 575 AA;

Query Match 100.0%; Score 2830; DB 5; Length 575;

Best Local Similarity 100.0%; Pred. No. 1.2e-149;

Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSOCVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 60  
Db 19 APAEPQGGSOCVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 78  
QY 61 GGVGRRRLWIGLQPLPPGCGDPRKLGPLRGFWVTGDNNTSYRWARLDLNGAPLCGLVC 120  
Db 79 GGVGRRRLWIGLQPLPPGCGDPRKLGPLRGFWVTGDNNTSYRWARLDLNGAPLCGLVC 138  
QY 121 AVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSIYGTTPFA 180  
Db 139 AVSAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSIYGTTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 258  
QY 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDQPGSYSCMCETGYRLAA 300  
Db 259 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDQPGSYSCMCETGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECYCNVNDLVGECVEPVPDPCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECYCNVNDLVGECVEPVPDPCFRANCEYQC 378  
QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQFCNOTACPADCDPNTQASCEPGEYIILDDGFI 420  
Db 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQFCNOTACPADCDPNTQASCEPGEYIILDDGFI 438  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSDSGSGEPPEPS 480  
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSDSGSGEPPEPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 14

AAE17528  
ID AAE17528 standard; protein; 557 AA.

XX AC AAE17528;

XX DT 22-APR-2002 (first entry)

XX DE Human thrombomodulin (TM) mutant, S474A.

XX KW Human; thrombomodulin; TM analogue; protein C; thrombotic disease;

antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor; TAFI; myocardial infarction; disseminated intravascular coagulation; DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina; acute respiratory distress syndrome; cancer; toxemia; septicemia; circulatory disorder; coronary embolism; pulmonary embolism; stroke; systemic coagulation disorder; immunosuppressive; cerebroprotective; pregnancy; anticoagulant; thrombolytic; venous occlusive condition; antibacterial; cardiant; mutant; mutein.

Homo sapiens.  
OS  
Synthetic.

Key Location/Qualifiers  
FT Misc-difference 474  
/note= "Wild-type Ser substituted with Ala; This location corresponds to position 492 of native thrombomodulin"

WO200198352-A2.  
27-DEC-2001.  
20-JUN-2001; 2001WO-US019590.  
21-JUN-2000; 2000US-0213678P.  
12-JUN-2001; 2001US-00880484.  
(SCHD ) SCHERING AG.  
Light D, Morser MJ, Nagashima M;  
WPI; 2002-154622/20.

Thrombomodulin analog with greater ability to potentiate thrombin-mediated activation of protein C and lesser ability to potentiate thrombin-mediated activation of thrombin-activatable fibrinolysis inhibitor for treating thrombotic disease.

Claim 11; Page; 53pp; English.

The invention relates to the design, production and use of thrombomodulin (TM) analogues that have the ability to enhance the thrombin-mediated activation of protein C which have a significantly reduced ability to promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI). TM analogue is useful for treating thrombotic disease in human, by administering it or a polynucleotide encoding it to the patient and useful for treating a thrombotic disease or condition in mammals. TM analogue is useful in antithrombotic therapy. TM analogue is also useful for treating diseases in which thrombus formation plays a significant etiological role including myocardial infarction, disseminated intravascular coagulation (DIC), deep vein thrombosis, septic shock, pulmonary embolism, acute respiratory distress syndrome, unstable angina and other arterial or venous occlusive conditions. TM analogue is also useful for treating various circulatory disorders including coronary or pulmonary embolism, stroke and systemic coagulation disorders including DIC which is often associated with septicemia, certain cancers and toxemia of pregnancy. The present sequence is human thrombomodulin mutant. Note: This sequence is not shown in the specification, but it is constructed based on human native thrombomodulin sequence SEQ.ID.NO.2 shown in fig 4 of the specification

Sequence 557 AA;

Query Match 99.9%; Score 2827; DB 5; Length 557;  
Best Local Similarity 99.8%; Pred. No. 1.7e-149;  
Matches 497; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCALYPGPATFELNASQICDGLRGLMTVRSSVAADVILLNGD 60  
DB 1 APAEPQGGSCQVEHDCALYPGPATFELNASQICDGLRGLMTVRSSVAADVILLNGD 60  
QY 61 GGVGRRRLWIGLQLPPGCGDKPKRLGFRQVWTDGNNNTSYSRWRLDLNGAPLCGLCV 120  
|||||

Db 61 GGVGRRRLWIGLQLPPGCGDKPKRLGFRQVWTDGNNNTSYSRWRLDLNGAPLCGLCV 120  
QY 121 AVSAEAATVPSEPIWEBOQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIITYGTFFA 180  
Db 121 AVSAEAATVPSEPIWEBOQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIITYGTFFA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
QY 241 IFGAPRCQCPAGAAALQADGRSCTASATOSCNDLCEHFCVNPDPQPGSYSCMCETGVRLAA 300  
Db 241 IFGAPRCQCPAGAAALQADGRSCTASATOSCNDLCEHFCVNPDPQPGSYSCMCETGVRLAA 300  
QY 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHYCNLYDLVDCGCEVPEVDPFCFRANCEYQC 360  
Db 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHYCNLYDLVDCGCEVPEVDPFCFRANCEYQC 360  
QY 361 QPLNNTSYLVCVCAEGFAPITPHEPHRCOMFCNCTACPADCDNTQASCCEPGYIILDDGFI 420  
Db 361 QPLNNTSYLVCVCAEGFAPITPHEPHRCOMFCNCTACPADCDNTQASCCEPGYIILDDGFI 420  
QY 421 CTDIDECENGFCGSGVCHNLPTFECICGPPDSALVRHIGTDCDSKVDGSGSGSEPPPS 480  
Db 421 CTDIDECENGFCGSGVCHNLPTFECICGPPDSALVRHIGTDCDSKVDGSGSGSEPPPS 480  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 481 PTPGSTLTTPPAVGLVHSG 498

RESULT 15  
AAE17525  
ID AAE17525 standard; protein; 557 AA.  
AC AAE17525;  
XX  
DT 22-APR-2002 (first entry)  
XX Human thrombomodulin (TM) mutant, M388L.  
DE Human; thrombomodulin; TM analogue; protein C; thrombotic disease;  
KW antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor;  
KW TAFI; myocardial infarction; disseminated intravascular coagulation; DIC;  
KW deep vein thrombosis; septic shock; pulmonary embolism; angina;  
KW acute respiratory distress syndrome; cancer; toxemia; septicemia;  
KW circulatory disorder; coronary embolism; pulmonary embolism; stroke;  
KW systemic coagulation disorder; immunosuppressive; cerebroprotective;  
KW pregnancy; anticoagulant; thrombolytic; venous occlusive condition;  
KW antibacterial; cardiant; mutant; mutein.  
XX Homo sapiens.  
OS Synthetic.  
OS  
Key Location/Qualifiers  
FT Misc-difference 388  
/note= "Wild-type Met substituted with Leu; This location corresponds to position 406 of native thrombomodulin"

WO200198352-A2.  
27-DEC-2001.  
20-JUN-2001; 2001WO-US019590.  
21-JUN-2000; 2000US-0213678P.  
12-JUN-2001; 2001US-00880484.  
(SCHD ) SCHERING AG.  
Light D, Morser MJ, Nagashima M;  
WPI; 2002-154622/20.





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:46:09 ; Search time 16.5 Seconds  
(without alignments)  
1558.166 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516

Perfect score: 2830

Sequence: 1 APAEPQGGSCQVEHDCFAL.....PSPTGSLTPPAVLVHSG 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description       |
|------------|--------|-------------|--------|-------|-------------------|
| 1          | 2830   | 100.0       | 498    | 2     | US-08-733-564-2   |
| 2          | 2830   | 100.0       | 575    | 1     | US-08-261-206A-59 |
| 3          | 2830   | 100.0       | 575    | 4     | US-09-880-484D-2  |
| 4          | 2826   | 99.9        | 575    | 1     | US-08-312-870-1   |
| 5          | 2826   | 99.9        | 575    | 6     | 5466668-6         |
| 6          | 2824   | 99.8        | 497    | 4     | US-09-331-793-4   |
| 7          | 2820   | 99.6        | 497    | 1     | US-08-312-870-3   |
| 8          | 2818   | 99.6        | 575    | 1     | US-08-170-290A-54 |
| 9          | 2765.5 | 97.4        | 572    | 6     | 5256770-7         |
| 10         | 2684   | 94.8        | 476    | 1     | US-08-014-723-1   |
| 11         | 2684   | 94.8        | 476    | 1     | US-08-110-011A-1  |
| 12         | 2684   | 94.8        | 494    | 1     | US-08-014-723-14  |
| 13         | 2684   | 94.8        | 494    | 1     | US-08-110-011A-14 |
| 14         | 2682   | 94.8        | 476    | 1     | US-08-014-723-2   |
| 15         | 2682   | 94.8        | 476    | 1     | US-08-014-723-18  |
| 16         | 2682   | 94.8        | 476    | 1     | US-08-110-011A-2  |
| 17         | 2682   | 94.8        | 476    | 1     | US-08-110-011A-18 |
| 18         | 2682   | 94.8        | 494    | 1     | US-08-014-723-16  |
| 19         | 2682   | 94.8        | 494    | 1     | US-08-110-011A-16 |
| 20         | 2604   | 92.0        | 475    | 1     | US-08-307-444A-1  |
| 21         | 2604   | 92.0        | 475    | 1     | US-08-587-389-1   |
| 22         | 2600   | 91.9        | 475    | 1     | US-08-587-389-2   |
| 23         | 2600   | 91.9        | 475    | 1     | US-08-587-389-3   |
| 24         | 2596   | 91.7        | 456    | 1     | US-08-307-444A-3  |
| 25         | 2596   | 91.7        | 456    | 1     | US-08-587-389-3   |
| 26         | 2592   | 91.6        | 456    | 1     | US-08-307-444A-4  |
| 27         | 2592   | 91.6        | 456    | 1     | US-08-587-389-4   |

28 2543 89.9 446 1 US-08-307-444A-5 Sequence 5, Appli  
29 2543 89.9 446 1 US-08-587-389-5 Sequence 5, Appli  
30 1617 57.1 275 1 US-08-312-870-7 Sequence 7, Appli  
31 1159 41.0 215 1 US-08-312-870-5 Sequence 5, Appli  
32 685 24.2 114 2 US-08-733-564-1 Sequence 1, Appli  
33 685 24.2 115 1 US-08-312-870-9 Sequence 9, Appli  
34 575 20.3 652 2 US-08-751-305-2 Sequence 2, Appli  
35 566.5 20.0 492 4 US-09-724-864-39 Sequence 39, Appli  
36 358 12.7 638 2 US-08-897-443-1 Sequence 1, Appli  
37 352 12.4 58 1 US-08-261-206A-3 Sequence 3, Appli  
38 350.5 12.4 915 4 US-09-907-794A-34 Sequence 34, Appli  
39 350.5 12.4 915 4 US-09-905-125A-34 Sequence 34, Appli  
40 350.5 12.4 915 4 US-09-902-775A-34 Sequence 34, Appli  
41 331.5 11.7 1964 4 US-08-467-997-1 Sequence 1, Appli  
42 328.5 11.6 956 2 US-08-897-443-3 Sequence 3, Appli  
43 323 11.4 1253 3 US-08-479-722B-4 Sequence 4, Appli  
44 320.5 11.3 2321 4 US-09-230-652-2 Sequence 2, Appli  
45 320 11.3 1394 6 5177197-30 Patent No. 5177197

## ALIGNMENTS

RESULT 1  
US-08-733-564-2  
; Sequence 2, Application US/08733564  
; Patent No. 5916874  
; GENERAL INFORMATION:  
; APPLICANT: FUJIWARA, Kenji  
; APPLICANT: MOCHIDA, Satoshi  
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/733,564  
; FILING DATE: 18 OCTOBER 1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SVENSSON, Leonard R.  
; REGISTRATION NUMBER: 30,330  
; REFERENCE/DOCKET NUMBER: 0216-0362P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 205-8000  
; TELEFAX: (703) 205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 498 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
US-08-733-564-2

Query Match 100.0%; Score 2830; DB 2; Length 498;  
Best Local Similarity 100.0%; Pred. No. 1e-192;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APAEPQGGSCQVEHDCFALYPGPATFNASQICDGLRGHLMTVRSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCQVEHDCFALYPGPATFNASQICDGLRGHLMTVRSVAADVLSLLNGD 60  
Qy 61 GGVRRLRWLQQLPPGCGDKRLGRLRGFWVTGDNNTSYSRWRLDLNGAPLCGLCV 120

Db 61 GGVGRRRLTGLQLPFGCGDPKELGFLRGFWTGDNTSYSRWARLDLNGAPLGLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPAGAAAASVITYTTPA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPAGAAAASVITYTTPA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMTCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQMTCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVFNPDQGSYSCMCEYRLAA 300  
Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVFNPDQGSYSCMCEYRLAA 300  
QY 301 DOHRCEDVDDCILEPSPQRCQRCVNTQGGFCHCYFNVDLVDGECVPEVDPFCFRANCEYQC 360  
Db 301 DOHRCEDVDDCILEPSPQRCQRCVNTQGGFCHCYFNVDLVDGECVPEVDPFCFRANCEYQC 360  
QY 361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDPNTQASCECEPGEYILDDGFI 420  
Db 361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDPNTQASCECEPGEYILDDGFI 420  
QY 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDALVRHIGTDCDVGKVDGSGSGEPPPS 480  
Db 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDALVRHIGTDCDVGKVDGSGSGEPPPS 480  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 481 PTPGSTLTTPPAVGLVHSG 498

## RESULT 2

US-08-261-206A-59  
; Sequence 59, Application US/08261206A  
; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Zushi, Mitichitaka  
; APPLICANT: Gomi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Birch, Stewart, Kolasch & Birch  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,206A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/740,492  
; FILING DATE: 03-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 216-275P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-241-1300  
; TELEFAX: 703-241-2848  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..575  
OTHER INFORMATION: /label= protein  
OTHER INFORMATION: /note= "human thrombomodulin"  
US-08-261-206A-59

Query Match 100.0%; Score 2830; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 1.2e-192; Mismatches 0; Indels 0; Gaps 0;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQPGSQCEVHDCFPALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60  
Db 19 APAEPQPGSQCEVHDCFPALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78  
QY 61 GGVGRRRLWIGLQLPFGCGDPKELGFLRGFWTGDNTSYSRWARLDLNGAPLGLCV 120  
Db 79 GGVGRRRLWIGLQLPFGCGDPKELGFLRGFWTGDNTSYSRWARLDLNGAPLGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPAGAAAASVITYTTPA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAVEPAGAAAASVITYTTPA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMTCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQMTCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVFNPDQGSYSCMCEYRLAA 300  
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVFNPDQGSYSCMCEYRLAA 318  
QY 301 DOHRCEDVDDCILEPSPQRCQRCVNTQGGFCHCYFNVDLVDGECVPEVDPFCFRANCEYQC 360  
Db 319 DOHRCEDVDDCILEPSPQRCQRCVNTQGGFCHCYFNVDLVDGECVPEVDPFCFRANCEYQC 378  
QY 361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDPNTQASCECEPGEYILDDGFI 420  
Db 379 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDPNTQASCECEPGEYILDDGFI 438  
QY 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDALVRHIGTDCDVGKVDGSGSGEPPPS 480  
Db 439 CTIDIDECENGFCGSGVCHNLPGTFECICGPDALVRHIGTDCDVGKVDGSGSGEPPPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

## RESULT 3

US-09-880-484D-2  
; Sequence 2, Application US/09880484D  
; Patent No. 6632791  
; GENERAL INFORMATION:  
; APPLICANT: Light, David  
; APPLICANT: Nagashima, Mariko  
; APPLICANT: Moiser, Michael J  
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use  
; FILE REFERENCE: 51863AUSM1  
; CURRENT APPLICATION NUMBER: US/09/880,484D  
; CURRENT FILING DATE: 2001-06-12  
; PRIOR APPLICATION NUMBER: 60/213,678  
; PRIOR FILING DATE: 2000-06-21  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

LENGTH: 575

TYPE: PRT

ORGANISM: Homo sapiens

US-09-880-484D-2

Query Match 100.0%; Score 2830; DB 4; Length 575;  
Best Local Similarity 100.0%; Pred. No. 1.2e-192;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFAALYGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 60  
DB 19 APAEPQGGSCQVEHDCFAALYGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 78  
QY 61 GGVGRRRLWIGLQLPPGCGDKRLGPRGFQVNTGDNNTSYSRWARLDLNGAPLCPLCV 120  
DB 79 GGVGRRRLWIGLQLPPGCGDKRLGPRGFQVNTGDNNTSYSRWARLDLNGAPLCPLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTTPA 180  
DB 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTTPA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAAREAPGAWDCSVENGCGCHACNA 240  
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAAREAPGAWDCSVENGCGCHACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDPFGSYSCMCETGYRLAA 300  
DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDPFGSYSCMCETGYRLAA 318  
QY 301 DOHRCEDVDDCILESPSPORCVNTQSGHCYNYDLVDGECVPEVPCFRANCETQC 360  
DB 319 DOHRCEDVDDCILESPSPORCVNTQSGHCYNYDLVDGECVPEVPCFRANCETQC 378  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECPGEGYLLDDGFI 420  
DB 379 QPLNQTSLVCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECPGEGYLLDDGFI 438  
QY 421 CTDIDECNGGFCGVCNHLNPTFECICGPDALVRHIGTDCSGKVDGSGSGGEPSPS 480  
DB 439 CTDIDECNGGFCGVCNHLNPTFECICGPDALVRHIGTDCSGKVDGSGSGGEPSPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
DB 499 PTPGSTLTTPPAVGLVHSG 516

#### RESULT 4

US-08-312-870-1

Sequence 1, Application US/08312870

Patent No. 5639625

GENERAL INFORMATION:

APPLICANT: Carlson, Craig W.

APPLICANT: Eason, Charles T.

TITLE OF INVENTION: Method for Detecting Antibodies to

TITLE OF INVENTION: Thrombomodulin in Patient

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESS: Richards, Medlock & Andrews

STREET: 1201 Elm Street, Suite 450

CITY: Dallas

STATE: Texas

COUNTRY: US

ZIP: 75270-2197

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/312,870

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hansen, Eugenia S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: OMRF B35150  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 214-939-4500

TELEFAX: 214-939-4600

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 575 amino acids

TYPE: aming acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Protein

LOCATION: 19..575

US-08-312-870-1

Query Match 99.9%; Score 2826; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 2.4e-192;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFAALYGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 60  
DB 19 APAEPQGGSCQVEHDCFAALYGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 78  
QY 61 GGVGRRRLWIGLQLPPGCGDKRLGPRGFQVNTGDNNTSYSRWARLDLNGAPLCPLCV 120  
DB 79 GGVGRRRLWIGLQLPPGCGDKRLGPRGFQVNTGDNNTSYSRWARLDLNGAPLCPLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTTPA 180  
DB 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTTPA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAAREAPGAWDCSVENGCGCHACNA 240  
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVOGHWAAREAPGAWDCSVENGCGCHACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDPFGSYSCMCETGYRLAA 300  
DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPNDPFGSYSCMCETGYRLAA 318  
QY 301 DOHRCEDVDDCILESPSPORCVNTQSGHCYNYDLVDGECVPEVPCFRANCETQC 360  
DB 319 DOHRCEDVDDCILESPSPORCVNTQSGHCYNYDLVDGECVPEVPCFRANCETQC 378  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECPGEGYLLDDGFI 420  
DB 379 QPLNQTSLVCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECPGEGYLLDDGFI 438  
QY 421 CTDIDECNGGFCGVCNHLNPTFECICGPDALVRHIGTDCSGKVDGSGSGGEPSPS 480  
DB 439 CTDIDECNGGFCGVCNHLNPTFECICGPDALVRHIGTDCSGKVDGSGSGGEPSPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
DB 499 PTPGSTLTTPPAVGLVHSG 516

#### RESULT 5

5466668-6

Patent No. 5466668

APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,

DAVID R.

TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR

PHARMACEUTICAL USE

NUMBER OF SEQUENCES: 57

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/155,346

FILING DATE: 22-NOV-1993

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 568,456
; FILING DATE: 15-AUG-1990
; APPLICATION NUMBER: 506,325
; FILING DATE: 09-APR-1990
; APPLICATION NUMBER: 406,941
; FILING DATE: 13-SEP-1989
; APPLICATION NUMBER: 345,374
; FILING DATE: 28-APR-1989
; SEQ ID NO:6;
; LENGTH: 575
; 546668-6

Query Match
Best Local Similarity 99.9%; Score 2826; DB 6; Length 575;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 78
QY 61 GGVRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120
Db 79 GGVRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPAAAAAASITYGTPEA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPAAAAAASITYGTPEA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 258
QY 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 259 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318
QY 301 DQHRCEVDVDCILEPSPQPCORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
Db 319 DQHRCEVDVDCILEPSPQPCORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 378
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPDNTOASCEPGEYIILDDGFI 420
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPDNTOASCEPGEYIILDDGFI 438
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKVGDGSGSGEPPPS 480
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKVGDGSGSGEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 6
US-09-331-793-4
; Sequence 4, Application US/09331793
; Patent No. 6500646
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi
; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/09/331,793
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-331-793-4

```

Query Match 99.8%; Score 2824; DB 4; Length 497;

```

Best Local Similarity 100.0%; Pred. No. 2.8e-192;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
Db 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
QY 61 GGVRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120
Db 61 GGVRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPAAAAAASITYGTPEA 180
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPAAAAAASITYGTPEA 180
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
QY 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 241 IFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
QY 301 DQHRCEVDVDCILEPSPQPCORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
Db 301 DQHRCEVDVDCILEPSPQPCORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPDNTOASCEPGEYIILDDGFI 420
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQACPDNTOASCEPGEYIILDDGFI 420
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKVGDGSGSGEPPPS 480
Db 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKVGDGSGSGEPPPS 480
QY 481 PTPGSTLTTPPAVGLVHSG 497
Db 481 PTPGSTLTTPPAVGLVHSG 497

```

## RESULT 7

```

US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carlson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600

```

```
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 497 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-312-870-3

Query Match          99.6%; Score 2820; DB 1; Length 497;
Best Local Similarity 99.8%; Pred. No. 5 3e-192;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
DB 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
QY 61 GGVGRRLWIGLQLPPGCGDKRLGRLGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
DB 61 GGVGRRLWIGLQLPPGCGDKRLGRLGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSTTYGTPFA 180
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSTTYGTPFA 180
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240
QY 241 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
DB 241 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
QY 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
DB 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECEGYIILDDGFI 420
DB 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECEGYIILDDGFI 420
QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPPEPS 480
DB 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPPEPS 480
QY 481 PTPGSTLTTPPAVGLVHVS 497
DB 481 PTPGSTLTTPPAVGLVHVS 497

RESULT 8
US-08-170-290A-54
/ Sequence 54, Application US/08170290A
/ Patent No. 5702931
/ GENERAL INFORMATION:
/ APPLICANT: Andrews, William H.
/ APPLICANT: Morser, Michael J.
/ APPLICANT: Zielesner, Laura R.
/ TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
/ NUMBER OF SEQUENCES: 63
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: James M. Heslin
/ STREET: 379 Lytton Ave.
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94301
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/08/170,290A
/ FILING DATE: 28-DEC-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US92/05573
/ FILING DATE: 01-JUL-1992
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/724,237
/ FILING DATE: 01-JUL-1991
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Heslin, James M.
/ REGISTRATION NUMBER: 29,541
/ REFERENCE/DOCKET NUMBER: 11972-58-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-326-2400
/ TELEFAX: 415-326-2422
/ INFORMATION FOR SEQ ID NO: 54:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 575 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-170-290A-54

Query Match          99.6%; Score 2818; DB 1; Length 575;
Best Local Similarity 99.6%; Pred. No. 8.7e-192;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
DB 19 APAEPQGGSCQVHDCSALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 78
QY 61 GGVGRRLWIGLQLPPGCGDKRLGRLGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
DB 79 GGVGRRLWIGLQLPPGCGDKRLGRLGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSTTYGTPFA 180
DB 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAVSTTYGTPFA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 258
QY 241 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
DB 259 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318
QY 301 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
DB 319 DQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 378
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECEGYIILDDGFI 420
DB 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECEGYIILDDGFI 438
QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPPEPS 480
DB 439 CTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPPEPS 498
QY 481 PTPGSTLTTPPAVGLVHVS 498
DB 499 PTPGSTLTTPPAVGLVHVS 516

RESULT 9
5256770-7
/ Patent No. 5256770
/ APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
/ DAVID R.
/ TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
```

```

; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7:
; LENGTH: 572
; 5256770-7

Query Match          97.4%; Score 2756.5; DB 6; Length 572;
Best Local Similarity 98.4%; Pred. No. 1.9e-187;
Matches 490; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
   |||||
Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 78
   |||||

QY 61 GGVGRRRLWIGLQPLPGCGDKPKLGRFQWVTGDNNTSYSRWARLDNGAPLCPLCV 120
   |||||
Db 79 GGVG-RRLWIGLQPLPGCGDKPKLGRFQWVTGDNNTSYSRWARLDNGAPLCPLCV 137
   |||||

QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTPEA 180
   |||||
Db 138 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTPEA 197
   |||||

QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCACNA 240
   |||||
Db 198 ARGADFOALPVGSSAAVAPLGLQLMCTA--GNVQGHWAREAPGAWDCSVENGCGEHCACNA 255
   |||||

QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSGSCMCTGYRLAA 300
   |||||
Db 256 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSGSCMCTGYRLAA 315
   |||||

QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDCGECVPEVDPGCFRANCEYQC 360
   |||||
Db 316 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDCGECVPEVDPGCFRANCEYQC 375
   |||||

QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCDPNTQASCECEGYILDDGFI 420
   |||||
Db 376 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCDPNTQASCECEGYILDDGFI 435
   |||||

QY 421 CTDIDCEGNGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDCKVGGDSDSGEPPPS 480
   |||||
Db 436 CTDIDCEGNGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDCKVGGDSDSGEPPPS 495
   |||||

QY 481 PTFGSTLTTPPAVLVHSG 498
   |||||
Db 496 PTFGSTLTTPPAVLVHSG 513
   |||||

```

## RESULT 10

```

US-08-014-723-1
; Sequence 1, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:

```

```

; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202

```

## COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248955 OPAT UR
; INFORMATION FOR SEQ ID NO: 1:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

US-08-014-723-1

```

```

Query Match          94.8%; Score 2684; DB 1; Length 476;
Best Local Similarity 99.4%; Pred. No. 2.1e-182;
Matches 473; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
   |||||
Db 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60
   |||||

QY 61 GGVGRRRLWIGLQPLPGCGDKPKLGRFQWVTGDNNTSYSRWARLDNGAPLCPLCV 120
   |||||
Db 61 GGVGRRRLWIGLQPLPGCGDKPKLGRFQWVTGDNNTSYSRWARLDNGAPLCPLCV 120
   |||||

QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTPEA 180
   |||||
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTPEA 180
   |||||

QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCACNA 240
   |||||
Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCACNA 240
   |||||

QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSGSCMCTGYRLAA 300
   |||||
Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSGSCMCTGYRLAA 300
   |||||

QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDCGECVPEVDPGCFRANCEYQC 360
   |||||
Db 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDCGECVPEVDPGCFRANCEYQC 360
   |||||

QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCDPNTQASCECEGYILDDGFI 420
   |||||
Db 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQNTACPADCDPNTQASCECEGYILDDGFI 420
   |||||

QY 421 CTDIDCEGNGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDCKVGGDSDSGE 476
   |||||
Db 421 CTDIDCEGNGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCDCKVGGDSDSGE 476
   |||||

```

## RESULT 11

```

US-08-110-011A-1
; Sequence 1, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:

```

```

; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

```

ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/110,011A  
FILING DATE: 23-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5354664man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)413-3000  
TELEFAX: (703)413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 476 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-110-011A-1

Query Match 94.8%; Score 2684; DB 1; Length 476;  
Best Local Similarity 99.4%; Pred. No. 2.1e-182;  
Matches 473; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 APAEPQPGSCQVCHDCFALYPGPATFLNASQICDGLRGHLMTVSSVAADVISLLNGD 60  
DB 1 APAEPQPGSCQVCHDCFALYPGPATFLNASQICDGLRGHLMTVSSVAADVISLLNGD 60  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYRWARLDLNGAPLCGLCV 120  
DB 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYRWARLDLNGAPLCGLCV 120  
QY 121 AVSAEATVPSEPIWEEOCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPA 180  
DB 121 AVSAEATVPSEPIWEEOCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
QY 241 IPGAPRCOCAGALQADGRSCTASATOSNDLCSEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 241 IPGAPRCOCAGALQADGRSCTASATOSNDLCSEHFCVNPDPQPGSYSCMCETGYRLAA 300  
QY 301 DOHRCEDVDDCILESPCPCORVNTQGGFECHCYNYDLVDGECVPEVDPFRANCEYQC 360  
DB 301 DOHRCEDVDDCILESPCPCORVNTQGGFECHCYNYDLVDGECVPEVDPFRANCEYQC 360  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECEGYLLDDGFI 420  
DB 361 QPLNQTSLYLCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECEGYLLDDGFI 420  
QY 421 CTIDCENGFCGSGVCHNLPGTFFCICGPDALSALVHIGTDCDSKVDGSGSGE 476  
DB 421 CTIDCENGFCGSGVCHNLPGTFFCICGPDALSALVHIGTDCDSKVDGSGSGE 476

RESULT 12

US-08-014-723-14

; Sequence 14, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi

APPLICANT: Iwasaki, Akio  
APPLICANT: Saino, Yushi  
APPLICANT: Kimura, Shigeru  
APPLICANT: Ohkuchi, Masao  
TITLE OF INVENTION: Thrombin-Binding Substance and Process  
TITLE OF INVENTION: For Preparing the Same  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/014,723  
FILING DATE: 19930208  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5273962man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)413-3000  
TELEFAX: (703)413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 494 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-014-723-14

Query Match 94.8%; Score 2684; DB 1; Length 494;  
Best Local Similarity 99.4%; Pred. No. 2.2e-182;  
Matches 473; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 APAEPQPGSCQVCHDCFALYPGPATFLNASQICDGLRGHLMTVSSVAADVISLLNGD 60  
DB 19 APAEPQPGSCQVCHDCFALYPGPATFLNASQICDGLRGHLMTVSSVAADVISLLNGD 78  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYRWARLDLNGAPLCGLCV 120  
DB 79 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYRWARLDLNGAPLCGLCV 138  
QY 121 AVSAEATVPSEPIWEEOCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPA 180  
DB 139 AVSAEATVPSEPIWEEOCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCHEACNA 258  
QY 241 IPGAPRCOCAGALQADGRSCTASATOSNDLCSEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 249 IPGAPRCOCAGALQADGRSCTASATOSNDLCSEHFCVNPDPQPGSYSCMCETGYRLAA 318  
QY 301 DOHRCEDVDDCILESPCPCORVNTQGGFECHCYNYDLVDGECVPEVDPFRANCEYQC 360  
DB 319 DOHRCEDVDDCILESPCPCORVNTQGGFECHCYNYDLVDGECVPEVDPFRANCEYQC 378  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECEGYLLDDGFI 420  
DB 379 QPLNQTSLYLCVCAEGFAPIPHEPRCOMFCNOTACPADCDPNTQASCECEGYLLDDGFI 438  
QY 421 CTIDCENGFCGSGVCHNLPGTFFCICGPDALSALVHIGTDCDSKVDGSGSGE 476



Db 439 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKGVDEYSGSE 494

## RESULT 13

US-08-110-011A-14  
; Sequence 14, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-14

Query Match 94.8%; Score 2684; DB 1; Length 494;  
Best Local Similarity 99.4%; Pred. No. 2.2e-182;  
Matches 473; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 APAEPQPGSGQVHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60  
Db 19 APAEPQPGSGQVHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 78  
QY 61 GGVRRRLWIGLQLPFGCGDPRKLGPRGFOWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
Db 79 GGVRRRLWIGLQLPFGCGDPRKLGPRGFOWVTGDNNTSYSRWARLDLNGAPLCGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAAAUSITGTFFA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAAAUSITGTFFA 198  
QY 181 ARGADFQALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGANDCSVENSGCHACNA 240  
Db 199 ARGADFQALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGANDCSVENSGCHACNA 258  
QY 241 IFGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSYCMCETGYRLAA 300  
Db 259 IFGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSYCMCETGYRLAA 318  
QY 301 DQHRCEDVDCCILEPSPQRCQVNTQGGFECHCYNYDLDVDCGECVEPVDPCFRANCEYQC 360

Db 319 DQHRCEDVDCCILEPSPQRCQVNTQGGFECHCYNYDLDVDCGECVEPVDPCFRANCEYQC 378  
QY 361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECPGYILLDDGFI 420  
Db 379 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECPGYILLDDGFI 438  
QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKGVDEYSGSE 476  
Db 439 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKGVDEYSGSE 494

## RESULT 14

US-08-014-723-2  
; Sequence 2, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESS: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-2

Query Match 94.8%; Score 2682; DB 1; Length 476;  
Best Local Similarity 99.2%; Pred. No. 3e-182;  
Matches 472; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 APAEPQPGSGQVHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60  
Db 1 APAEPQPGSGQVHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60  
QY 61 GGVRRRLWIGLQLPFGCGDPRKLGPRGFOWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
Db 61 GGVRRRLWIGLQLPFGCGDPRKLGPRGFOWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAAAUSITGTFFA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAVEPGAAAAAUSITGTFFA 180



QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
 Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 QY 301 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDCGCEVPEVDPDFCFRANCEYQC 360  
 Db 301 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDCGCEVPEVDPDFCFRANCEYQC 360  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCDPNTQASCECEGYIILDDGFI 420  
 Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCDPNTQASCECEGYIILDDGFI 420  
 QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGS 476  
 Db 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGS 476

## RESULT 15

US-08-014-723-18

; Sequence 19, Application US/08014723

; Patent No. 5273962

## ; GENERAL INFORMATION:

; APPLICANT: Doi, Takeshi

; APPLICANT: Iwasaki, Akio

; APPLICANT: Saino, Yushi

; APPLICANT: Kimura, Shigeru

; APPLICANT: Okuchi, Masao

; TITLE OF INVENTION: Thrombin-Binding Substance and Process

; TITLE OF INVENTION: For Preparing the Same

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER &amp; NEUSTADT,

; ADDRESS: P.C.

; STREET: 1755 Jefferson Davis Highway, Fourth Floor

; CITY: Arlington

; STATE: Virginia

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/014,723

; FILING DATE: 19930208

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Oblon, No. 5273962man F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 80-071-0 CIP

; TELEPHONE: (703)413-3000

; TELEFAX: (703)413-2220

; TELEFAX: (703)413-2220

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 476 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 469

; OTHER INFORMATION: /note= "acidic amino acid"

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 470

; OTHER INFORMATION: /note= "acidic amino acid"

; US-08-014-723-18

Query Match 94.8%; Score 2682; DB 1; Length 476;  
 Best Local Similarity 99.4%; Pred. No. 3e-182;  
 Matches 472; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSILLNGD 60  
 Db 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSILLNGD 60  
 QY 61 GGVGRRRLWIGLQLPFGCGDKPRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120  
 Db 61 GGVGRRRLWIGLQLPFGCGDKPRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFPATCEPLAVEPAAAAAASITTYGTFFA 180  
 Db 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFPATCEPLAVEPAAAAAASITTYGTFFA 180  
 QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
 Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 QY 301 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDCGCEVPEVDPDFCFRANCEYQC 360  
 Db 301 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDCGCEVPEVDPDFCFRANCEYQC 360  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCDPNTQASCECEGYIILDDGFI 420  
 Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCNQTACPADCDPNTQASCECEGYIILDDGFI 420  
 QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGS 475  
 Db 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDCDSKVDGSGSGS 475

Search completed: June 9, 2004, 08:51:29

Job time : 17.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:50:04 ; Search time 42 Seconds  
(without alignments)  
3335.873 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516

Perfect score: 2830

Sequence: 1 APAEPQGGSCQVEHDCFAL.....PSPTGSLTPPAVLGHVSHG 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
17: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
18: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 2830  | 100.0       | 575    | 9     | US-09-938-405-2     |
| 2          | 2830  | 100.0       | 575    | 10    | US-09-860-464-2     |
| 3          | 2830  | 100.0       | 575    | 14    | US-10-438-648-2     |
| 4          | 2830  | 100.0       | 575    | 15    | US-10-410-195-2     |
| 5          | 2826  | 99.9        | 575    | 14    | US-10-150-440-3     |
| 6          | 2826  | 99.9        | 575    | 15    | US-10-373-801-29    |
| 7          | 2826  | 99.9        | 631    | 14    | US-10-150-440-1     |
| 8          | 2824  | 99.8        | 497    | 15    | US-10-298-786-4     |
| 9          | 2815  | 99.5        | 575    | 15    | US-10-094-886-196   |
| 10         | 1036  | 36.6        | 239    | 15    | US-10-104-047-2759  |
| 11         | 731   | 25.8        | 418    | 12    | US-10-427-805-2     |
| 12         | 711   | 25.1        | 397    | 12    | US-10-427-805-3     |
| 13         | 578   | 20.4        | 645    | 14    | US-10-029-386-33151 |
| 14         | 578   | 20.4        | 652    | 16    | US-10-408-765A-1422 |
| 15         | 576   | 20.4        | 652    | 9     | US-09-789-919-96    |

|    |       |      |      |    |                     |                    |
|----|-------|------|------|----|---------------------|--------------------|
| 16 | 575   | 20.3 | 652  | 12 | US-10-211-462-131   | Sequence 131, App  |
| 17 | 575   | 20.3 | 652  | 14 | US-10-021-660-83    | Sequence 83, Appl  |
| 18 | 574   | 20.3 | 467  | 12 | US-10-210-172-176   | Sequence 176, App  |
| 19 | 574   | 20.3 | 757  | 12 | US-09-918-715-177   | Sequence 177, App  |
| 20 | 574   | 20.3 | 757  | 12 | US-09-918-715-196   | Sequence 196, App  |
| 21 | 574   | 20.3 | 757  | 15 | US-10-262-445-107   | Sequence 107, App  |
| 22 | 566.5 | 20.0 | 765  | 12 | US-09-918-715-190   | Sequence 190, App  |
| 23 | 566.5 | 20.0 | 765  | 12 | US-09-918-715-291   | Sequence 291, App  |
| 24 | 538   | 19.0 | 644  | 9  | US-09-789-919-62    | Sequence 62, Appl  |
| 25 | 538   | 19.0 | 644  | 9  | US-09-789-919-73    | Sequence 73, Appl  |
| 26 | 414   | 14.6 | 1574 | 12 | US-09-825-751A-77   | Sequence 77, Appl  |
| 27 | 374   | 13.2 | 2906 | 15 | US-10-015-115-60    | Sequence 60, Appl  |
| 28 | 368   | 13.0 | 1231 | 12 | US-10-363-616-482   | Sequence 482, App  |
| 29 | 367.5 | 13.0 | 1051 | 16 | US-10-309-290-10    | Sequence 10, Appl  |
| 30 | 367.5 | 13.0 | 1131 | 16 | US-10-309-290-8     | Sequence 8, Appl   |
| 31 | 367.5 | 13.0 | 1145 | 16 | US-10-309-290-12    | Sequence 12, Appl  |
| 32 | 367.5 | 13.0 | 1184 | 16 | US-10-309-290-6     | Sequence 6, Appl   |
| 33 | 367   | 13.0 | 2911 | 15 | US-10-015-115-59    | Sequence 59, Appl  |
| 34 | 365   | 12.9 | 2871 | 12 | US-09-825-751A-68   | Sequence 68, Appl  |
| 35 | 365   | 12.9 | 2911 | 15 | US-10-298-027-162   | Sequence 162, App  |
| 36 | 365   | 12.9 | 2911 | 16 | US-10-408-765A-421  | Sequence 421, App  |
| 37 | 363.5 | 12.8 | 2695 | 15 | US-10-015-115-14    | Sequence 14, Appl  |
| 38 | 363   | 12.8 | 1393 | 12 | US-10-312-352-21    | Sequence 21, Appl  |
| 39 | 363   | 12.8 | 2871 | 15 | US-10-015-115-57    | Sequence 57, Appl  |
| 40 | 363   | 12.8 | 3002 | 15 | US-10-015-115-56    | Sequence 56, Appl  |
| 41 | 363   | 12.8 | 3002 | 16 | US-10-408-765A-195  | Sequence 195, App  |
| 42 | 359.5 | 12.7 | 1218 | 15 | US-10-161-493-102   | Sequence 102, App  |
| 43 | 359   | 12.7 | 741  | 16 | US-10-408-765A-1187 | Sequence 1187, App |
| 44 | 359   | 12.7 | 782  | 12 | US-10-428-275-176   | Sequence 176, App  |
| 45 | 359   | 12.7 | 794  | 11 | US-09-833-245-1292  | Sequence 1292, Ap  |

#### ALIGNMENTS

#### RESULT 1

US-09-938-405-2  
; Sequence 2, Application US/09938405  
; Patent No. US20020111296A1  
; GENERAL INFORMATION:  
; APPLICANT: Festoff, Barry W.  
; APPLICANT: Morser, Michael J.  
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury  
; FILE REFERENCE: 51960AUSM1  
; CURRENT APPLICATION NUMBER: US/09/938,405  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: 60/229,714  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-938-405-2

|                       |         |  |           |            |    |        |      |
|-----------------------|---------|--|-----------|------------|----|--------|------|
| Query Match           | 100.0%; | Score  | 2830;     | DB         | 9; | Length | 575; |
| Best Local Similarity | 100.0%; | Pred. No.  | 5.7e-194; |            |    |        |      |
| Matches               | 498;    | Conservative   | 0;        | Mismatches | 0; | Indels | 0;   |
| Gaps                  | 0;      |  |           |            |    |        |      |
| QY                    | 1       | APAEPQGGSCQVEHDCFALYFGPATFNASQICDLRGHLMTVRSSVAADVISLLNGD     | 60        |            |    |        |      |
| Db                    | 19      | APAEPQGGSCQVEHDCFALYFGPATFNASQICDLRGHLMTVRSSVAADVISLLNGD     | 78        |            |    |        |      |
| QY                    | 61      | GGVGRRLWIGLQPPGCGDKPKLGRFGQWVTGDNNTSYSRWARLDNGAPLCPLCV       | 120       |            |    |        |      |
| Db                    | 79      | GGVGRRLWIGLQPPGCGDKPKLGRFGQWVTGDNNTSYSRWARLDNGAPLCPLCV       | 138       |            |    |        |      |
| QY                    | 121     | AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFHFPATCEPLAVEPGAAAAAIVTGTFFA | 180       |            |    |        |      |
| Db                    | 139     | AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFHFPATCEPLAVEPGAAAAAIVTGTFFA | 198       |            |    |        |      |
| QY                    | 181     | ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCHEACNA | 240       |            |    |        |      |

Db 199 ARGADFOALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 360  
Db 319 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 378  
QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
QY 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 480  
Db 439 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516  
RESULT 2  
US-09-880-464-2  
; Sequence 2, Application US/09880464  
; Publication No. US20030011623A1  
; GENERAL INFORMATION:  
; APPLICANT: Light, David  
; APPLICANT: Morser, Michael J  
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use  
; FILE REFERENCE: Improved Thrombomodulin Analogs  
; CURRENT APPLICATION NUMBER: US/09/880,464  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/213,678  
; PRIOR FILING DATE: 2000-06-21  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-464-2

Query Match 100.0%; Score 2830; DB 10; Length 575;  
Best Local Similarity 100.0%; Pred. No. 5.7e-194;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60  
Db 19 APAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 78  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCPGLCV 120  
Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCPGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAVSITYGTPTA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAVSITYGTPTA 198  
QY 181 ARGADFOALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 199 ARGADFOALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 360  
Db 319 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 378

QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
QY 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 480  
Db 439 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516  
RESULT 3  
US-10-438-648-2  
; Sequence 2, Application US/10438648  
; Publication No. US20030186883A1  
; GENERAL INFORMATION:  
; APPLICANT: Light, David  
; APPLICANT: Nagashima, Mariko  
; APPLICANT: Morser, Michael J  
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use  
; FILE REFERENCE: 51863AUSD1  
; CURRENT APPLICATION NUMBER: US/10/438,648  
; CURRENT FILING DATE: 2003-05-14  
; PRIOR APPLICATION NUMBER: US 60/213,678  
; PRIOR FILING DATE: 2000-06-21  
; PRIOR APPLICATION NUMBER: US 09/880,484  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-438-648-2

Query Match 100.0%; Score 2830; DB 14; Length 575;  
Best Local Similarity 100.0%; Pred. No. 5.7e-194;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60  
Db 19 APAEPQPGSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 78  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCPGLCV 120  
Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCPGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAVSITYGTPTA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAVSITYGTPTA 198  
QY 181 ARGADFOALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 199 ARGADFOALPVGSSAAVPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSQCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 360  
Db 319 DQHRCEVDVDCILLEPSPCPORCVNTQGGFECHCYPNYDLVDGECVBPVDPCEFRANCEYOC 378  
QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
QY 421 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 480  
Db 439 CTIDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDCKGVDGSGSGSEPPPS 498

```
Qy 481 PTPGSLTPPAVLVHSG 498
Db 499 PTPGSLTPPAVLVHSG 516

RESULT 4
US-10-410-195-2
; Sequence 2, Application US/10410195
; Publication No. US20040002446A1
; GENERAL INFORMATION:
; APPLICANT: Pestoff, Barry W.
; APPLICANT: Morser, Michael J.
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury
; FILE REFERENCE: 51960AUSM1
; CURRENT APPLICATION NUMBER: US/10/410,195
; CURRENT FILING DATE: 2003-08-10
; PRIOR APPLICATION NUMBER: US/09/938,405
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/229,714
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-195-2

Query Match 100.0%; Score 2830; DB 15; Length 575;
Best Local Similarity 100.0%; Pred. No. 5,7e-194;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 APAEPQGGSCVCHDCFCALYPGPATFLNASQICDGLRGHLMTYRSSVAADVISILLNGD 60
Db 19 APAEPQGGSCVCHDCFCALYPGPATFLNASQICDGLRGHLMTYRSSVAADVISILLNGD 78
Qy 61 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQVMTGDNNTSYRWARLDLNGAPLCGPLCV 120
Db 79 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQVMTGDNNTSYRWARLDLNGAPLCGPLCV 138
Qy 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTFFA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTFFA 198
Qy 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCNA 258
Qy 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318
Qy 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYCNVNTQACDPCDNTQASCEPEGYILDDGFI 360
Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYCNVNTQACDPCDNTQASCEPEGYILDDGFI 378
Qy 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCTACPDNTQASCEPEGYILDDGFI 420
Db 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCTACPDNTQASCEPEGYILDDGFI 438
Qy 421 CTDIDECENGFCGVCVCHNLPTFECICGPDSSALVRHIGTDCDCKVDDGSGSGSEPPPS 480
Db 439 CTDIDECENGFCGVCVCHNLPTFECICGPDSSALVRHIGTDCDCKVDDGSGSGSEPPPS 498
Qy 481 PTPGSLTPPAVLVHSG 498
Db 499 PTPGSLTPPAVLVHSG 516

RESULT 5
US-10-150-440-3
; Sequence 3, Application US/10150440
```

```
; Publication No. US20030022249A1
; GENERAL INFORMATION:
; APPLICANT: Schmitz, Juergen
; APPLICANT: Dzionek, Andrzej
; APPLICANT: Buck, David William
; TITLE OF INVENTION: ANTIGEN-BINDING FRAGMENTS THAT RECOGNIZE
; TITLE OF INVENTION: A SUBSET OF DENDRITIC CELLS AND METHODS OF USE THEREOF
; FILE REFERENCE: 212302001200
; CURRENT APPLICATION NUMBER: US/10/150,440
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/714,712
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/291,561
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/197,205
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 60/196,824
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/180,775
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: US 60/179,003
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/167,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/165,555
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-150-440-3

Query Match 99.9%; Score 2826; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.1e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 APAEPQGGSCVCHDCFCALYPGPATFLNASQICDGLRGHLMTYRSSVAADVISILLNGD 60
Db 19 APAEPQGGSCVCHDCFCALYPGPATFLNASQICDGLRGHLMTYRSSVAADVISILLNGD 78
Qy 61 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQVMTGDNNTSYRWARLDLNGAPLCGPLCV 120
Db 79 GVGRRRLWIGLQPLPGCGDPKRLGRLGFGQVMTGDNNTSYRWARLDLNGAPLCGPLCV 138
Qy 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTFFA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIYGTFFA 198
Qy 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHCNA 258
Qy 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318
Qy 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYCNVNTQACDPCDNTQASCEPEGYILDDGFI 360
Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYCNVNTQACDPCDNTQASCEPEGYILDDGFI 378
Qy 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCTACPDNTQASCEPEGYILDDGFI 420
Db 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCTACPDNTQASCEPEGYILDDGFI 438
Qy 421 CTDIDECENGFCGVCVCHNLPTFECICGPDSSALVRHIGTDCDCKVDDGSGSGSEPPPS 480
Db 439 CTDIDECENGFCGVCVCHNLPTFECICGPDSSALVRHIGTDCDCKVDDGSGSGSEPPPS 498
Qy 481 PTPGSLTPPAVLVHSG 498
Db 499 PTPGSLTPPAVLVHSG 516
```

```

RESULT 6
US-10-373-801-29
; Sequence 29, Application US/10373801
; Publication No. US20040005644A1
; GENERAL INFORMATION:
; APPLICANT: Yibai Pharmaceutical (USA)
; TITLE OF INVENTION: Method and composition for detection and treatment of breast cancer
; FILE REFERENCE: 12399-00
; CURRENT APPLICATION NUMBER: US/10/373,801
; CURRENT FILING DATE: 2003-02-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-373-801-29

Query Match          99.9%; Score 2826; DB 15; Length 575;
Best Local Similarity 99.8%; Pred. No. 1.1e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60
Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78
QY 61 GGVRRRLWIGLQIPPGCGDKPLGLRGPQWVTGDNNTSYSRWARLDNGAPLGCPLCV 120
Db 79 GGVRRRLWIGLQIPPGCGDKPLGLRGPQWVTGDNNTSYSRWARLDNGAPLGCPLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAVSITYGTTPA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAVSITYGTTPA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 258
QY 241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 300
Db 259 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 318
QY 301 DQHRCEVDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFRANCEYQC 360
Db 319 DQHRCEVDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFRANCEYQC 378
QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 420
Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 438
QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVGDSGSGSEPPPS 480
Db 439 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVGDSGSGSEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 7
US-10-150-440-1
; Sequence 1, Application US/10150440
; Publication No. US20030022249A1
; GENERAL INFORMATION:
; APPLICANT: Schmitz, Juergen
; APPLICANT: Dzionek, Andrzej
; APPLICANT: Buck, David William
; TITLE OF INVENTION: ANTIGEN-BINDING FRAGMENTS THAT RECOGNIZE
; TITLE OF INVENTION: A SUBSET OF DENDRITIC CELLS AND METHODS OF USE THEREOF
; FILE REFERENCE: 212302001200
; CURRENT APPLICATION NUMBER: US/10/150,440

```

```

; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/714,712
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: US 60/291,561
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/197,205
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 60/196,824
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/180,775
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: US 60/179,003
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/167,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/165,555
; PRIOR FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-150-440-1

Query Match          99.9%; Score 2826; DB 14; Length 631;
Best Local Similarity 99.8%; Pred. No. 1.2e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60
Db 75 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 134
QY 61 GGVRRRLWIGLQIPPGCGDKPLGLRGPQWVTGDNNTSYSRWARLDNGAPLGCPLCV 120
Db 135 GGVRRRLWIGLQIPPGCGDKPLGLRGPQWVTGDNNTSYSRWARLDNGAPLGCPLCV 194
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAVSITYGTTPA 180
Db 195 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAVSITYGTTPA 254
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 240
Db 255 ARGADFOALPVGSSAAVAPLGLQIMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 314
QY 241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 300
Db 315 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 374
QY 301 DQHRCEVDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFRANCEYQC 360
Db 375 DQHRCEVDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPFRANCEYQC 434
QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 420
Db 435 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 494
QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVGDSGSGSEPPPS 480
Db 495 CTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVGDSGSGSEPPPS 554
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 555 PTPGSTLTTPPAVGLVHSG 572

RESULT 8
US-10-298-796-4
; Sequence 4, Application US/10298796
; Publication No. US20030220490A1
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi

```

; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS  
; FILE REFERENCE: 1110-253P  
; CURRENT APPLICATION NUMBER: US/10/298,796  
; CURRENT FILING DATE: 2002-11-19  
; PRIOR APPLICATION NUMBER: US/09/331,793  
; PRIOR FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 497  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-298-796-4

Query Match 99.8%; Score 2824; DB 15; Length 497;  
Best Local Similarity 100.0%; Pred. No. 1.3e-193; Indels 0; Gaps 0;  
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCQVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60  
QY 61 GGVRRLWIGLQPPGCGDPKRLGPRGFQWVTGDNNTSYSRWARLDLNGAPLPGPLCV 120  
Db 61 GGVRRLWIGLQPPGCGDPKRLGPRGFQWVTGDNNTSYSRWARLDLNGAPLPGPLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSCMCTGYRLAA 300  
Db 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSCMCTGYRLAA 300  
QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDPFCFRANCEYOC 360  
Db 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDPFCFRANCEYOC 360  
QY 361 QPLNQTSLYLCVCAEGFAPLPHBPHRCMPFNQACPADCPNTQASCECEGVIYLDGFI 420  
Db 361 QPLNQTSLYLCVCAEGFAPLPHBPHRCMPFNQACPADCPNTQASCECEGVIYLDGFI 420  
QY 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPFPPS 480  
Db 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALVRHIGTDCDSKGVGDGSGSGEPFPPS 480  
QY 481 PTPGSTLTTPPAVGLVHS 497  
Db 481 PTPGSTLTTPPAVGLVHS 497

RESULT 9  
US-10-094-886-196  
; Publication 196, Application US/10094886  
; Publication No. US20040002120A1  
; GENERAL INFORMATION:  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Vernet, Corine A.  
; APPLICANT: Li, Li  
; APPLICANT: Gorman, Linda  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shenoy, Suresh

; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Taupier, Raymond J., Jr.  
; APPLICANT: Miller, Charles  
; APPLICANT: Casman, Stacie  
; APPLICANT: Pena, Carol  
; APPLICANT: Gangolli, Bsha  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Pochart, Pascal  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Laroche, William  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-290 B  
; CURRENT APPLICATION NUMBER: US/10/094,886  
; CURRENT FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: 60/274,322  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/313,182  
; PRIOR FILING DATE: 2001-08-17  
; PRIOR APPLICATION NUMBER: 60/288,052  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: 60/318,510  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/274,281  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/314,018  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,849  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/296,693  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: 60/313,626  
; PRIOR FILING DATE: 2001-08-21  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 298  
; SOFTWARE: PatentIn 2.1  
; SEQ ID NO 196  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-886-196

Query Match 99.5%; Score 2815; DB 15; Length 575;  
Best Local Similarity 99.4%; Pred. No. 6.7e-193;  
Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 60  
Db 19 APAEPQGGSCQVHEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVLSLLNGD 78  
QY 61 GGVRRLWIGLQPPGCGDPKRLGPRGFQWVTGDNNTSYSRWARLDLNGAPLPGPLCV 120  
Db 79 GGVRRLWIGLQPPGCGDPKRLGPRGFQWVTGDNNTSYSRWARLDLNGAPLPGPLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPQSGYSCMCTGYRLAA 318

QY 301 DOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDPFCFRANCEYQC 360  
DB 319 DQHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDPFCFRANCEYQC 378  
QY 361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPNTQASCECEGYLDGFI 420  
DB 379 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPNTQASCECEGYLDGFI 438  
QY 421 CTDIDCEGCGFCGCHNLPGTFECICGPDSDALVRHIGTDCDSKGVGDGSGSGEPSPS 480  
DB 439 CTDIDCEGCGFCGCHNLPGTFECICGPDSDALVRHIGTDCDSKGVGDGSGSGEPSPS 498  
QY 481 PTPGSTLTTPPAVLVHSG 498  
DB 499 PTPGSTLTTPPAVLVHSG 516

## RESULT 10

US-10-104-047-2759  
; Sequence 2759, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2759  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-2759

Query Match 36.6%; Score 1036; DB 15; Length 239;  
Best Local Similarity 86.5%; Pred. No. 3.1e-66;  
Matches 180; Conservative 0; Mismatches 0; Indels 28; Gaps 1;  
QY 291 MCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDP 350  
DB 1 MCETGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGCEVPVDP 60  
QY 351 CFRANCEYQCQPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPNTQASCECP 410  
DB 61 CFRANCEYQCQPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNQTACPADCDPNTQASCE-- 118  
QY 411 EGYLDGFICTDIDCEGCGFCGCHNLPGTFECICGPDSDALVRHIGTDCDSKGVGDG 470  
DB 119 -----CHNLPGTFECICGPDSDALVRHIGTDCDSKGVGDG 152  
QY 471 DSGSGEPSPPTPGSTLTTPPAVLVHSG 498  
DB 153 DSGSGEPSPPTPGSTLTTPPAVLVHSG 180

## RESULT 11

US-10-427-805-2  
; Sequence 2, Application US/10427805  
; Publication No. US20040063632A1  
; GENERAL INFORMATION:  
; APPLICANT: McLean, Kirk  
; TITLE OF INVENTION: No. US20040063632A1 Tissue Factor Targeted Thrombomodulin Fusio  
; FILE REFERENCE: 52295AUSM1  
; CURRENT APPLICATION NUMBER: US/10/427,805  
; PRIOR FILING DATE: 2003-04-30  
; PRIOR APPLICATION NUMBER: US 60/376,566  
; PRIOR FILING DATE: 2002-05-01

; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 418  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence of scFv(TF)3e10-TM1456 fusion protein  
US-10-427-805-2

Query Match 25.1%; Score 711; DB 12; Length 418;  
Best Local Similarity 50.2%; Pred. No. 3.6e-44;  
Matches 148; Conservative 18; Mismatches 56; Indels 73; Gaps 8;  
QY 243 GAPRCQCPAGALOAD-GRSCTASATQSCNDLCEHFVCPNPDPQGSYSCMCETGYRLAAD 301  
DB 150 GAPNFMLTQPHSVASPGKTVITISCTRSSGSVASYVQWYQORPGSSPTTV-----IYED 204  
QY 302 QHRCEDVDDCILEPSPQRCVNT-----CVNT-----OGGFECHCYPNYDLV-- 340  
DB 205 NHR-----PSGVDPDRFSGSIDTSSNSASLTISGLKTEDEADYICQSDSNLVVF 254  
QY 341 -----DGEVPEVDPFCFRANCEYQCQPLNQTSYLVCVCAEGF 376  
DB 255 GGGTKLTVLGAAGGGGGGGGGGGVPEVDPFCFRANCEYQCQPLNQTSYLVCVCAEGF 314  
QY 377 APIPEPHRCQMFNQTACPADCDPNTQASCECEPYLDDGFICTDIDCEGCGFCGSGV 436  
DB 315 APIPEPHRCQMFNQTACPADCDPNTQASCECEPYLDDGFICTDIDCEGCGFCGSGV 374  
QY 437 CHNLPGTFECICGPDSDALVRHIGTDCDSKGVGDGSGSGEPSPPTPGSTLTTPPA 491  
DB 375 CHNLPGTFECICGPDSDALAGQIGTDC-----AAAGAPVYPDP---LEPRA 417

## RESULT 12

US-10-427-805-3  
; Sequence 3, Application US/10427805  
; Publication No. US20040063632A1  
; GENERAL INFORMATION:  
; APPLICANT: Light, David  
; APPLICANT: McLean, Kirk  
; TITLE OF INVENTION: No. US20040063632A1 Tissue Factor Targeted Thrombomodulin Fusio  
; FILE REFERENCE: 52295AUSM1  
; CURRENT APPLICATION NUMBER: US/10/427,805  
; PRIOR FILING DATE: 2003-04-30  
; PRIOR APPLICATION NUMBER: US 60/376,566  
; PRIOR FILING DATE: 2002-05-01  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Amino acid sequence of scFv(TF)3e10-TM1456delta fusio  
US-10-427-805-3

Query Match 25.1%; Score 711; DB 12; Length 397;  
Best Local Similarity 54.6%; Pred. No. 9.1e-43;  
Matches 136; Conservative 14; Mismatches 39; Indels 60; Gaps 5;  
QY 259 GRSCTASATQSCNDLCEHFVCPNPDPQGSYSCMCETGYRLAADQHRCEDVDDCILEPSPC 318  
DB 164 GKTVITISCTRSSGSVASYVQWYQORPGSSPTTV-----IYEDNHR-----PSGV 208  
QY 319 POR---CVNT-----OGGFECHCYPNYDLV----- 340  
DB 209 PDRFSGSIDTSSNSASLTISGLKTEDEADYICQSDSNLVVFGGTTKLTVLGAAGGGG 268  
QY 341 -----DGEVPEVDPFCFRANCEYQCQPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNQT 393

Db 269 SGGGGGGGSGVDFVDFCFRANCEYQCOPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQT 328  
QY 394 ACADCDPNTQASCECEGYLLDDGFTCTDIDSCENGCGFCGCHNLPGTFECICGPDSA 453  
Db 329 ACADCDPNTQASCECEGYLLDDGFTCTDIDSCENGCGFCGCHNLPGTFECICGPDSA 388  
QY 454 LVRHIGTDC 462  
Db 389 LAGQIGTDC 397

RESULT 13  
US-10-029-386-33151  
; Sequence 33151, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: ABOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 33151  
; LENGTH: 645  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO ALL18508.6  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.83  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
; OTHER INFORMATION: SWISSPROT HIT: P07204, EVALUATION 7.00e-44  
US-10-029-386-33151

Query Match 20.4%; Score 578; DB 14; Length 645;  
Best Local Similarity 31.6%; Pred. No. 4.9e-33;  
Matches 175; Conservative 52; Mismatches 201; Indels 126; Gaps 31;

QY 6 QPGGSG-----CWEHDFALYPGATFLNASQICDGLRGLMTVRSSVAAD-----VI 53  
Db 17 QPGAGTGADTEAVVCGTACTYTAHSGKLSAAEAQNHCNQGNGNLATVKSKEEAHQVRVL 76  
QY 54 SLLNGDGGVGR--RLWIGLQLP--CGDPKRLGLRFGQWVTGNNTSYSRWALDLN 110  
Db 77 AQLRREALTARMKFWIGLQREKGLDPSL--PLKGFVWVGGEPTPYSNWHKELRN 134  
QY 111 GAPLCGPLCVA--VSAAEATVPSE--PIWEEQOC-----EVKADGFLCFEHPATCRPLA 161  
Db 135 SC--ISKRCVSLLLDLSQPLLPRLPKWSEGPCGSPGSGNIEGFVCFKFGMCRPLA 192  
QY 162 V-EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMC--TAPPGAQVQHW-- 217  
Db 193 LGPG-----QVYTTPTFTTSSLEAVPFAAANVA-----CGEGDKDETQSHYFL 239  
QY 218 -AREAPGAWD-----CSVENGGCEHAC--NAIPGAPRCQCPAGALQADGRS 261  
Db 240 CKEAPADVDMWSSGPLCVSPKYGKFNNGGCHQDCFEFGDGSFLCGCRPFGLLDLVT 299  
QY 262 CTASATQSCNDLCE--HFCVNPDPQSGYSKCMETGYRLAADQHRCEDVDDCILEPSPCP 319  
Db 300 C-ASRNPSSSPCRGGATCVLGP-HGKNYTCRCPOGYQLDSSQLDCVDVDEC--QDSPCA 355  
QY 320 QRCVNTQGGFECHCYNDYLDVGECEVPVDFPCFRANCEYQCOPLNQTSYLVCVCAEGFAP 379  
Db 356 QECVNTPGGFRCECWGY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395  
QY 380 PHEPHRCQMFNCQACPADCDPNTQAS--CECEGYIL--DDGFICTDIDEC--ENGGFC 433  
Db 396 -----CAQGC--TNTDGSFHCSEEGYVLAGEDGTQCQDVDCVGGGGLC 439

QY 434 SGVCHNLPGTFECICGPDGALVRHIGTDCDSGKV-----DGGDSGSGE-----P 477  
Db 440 DSLCFTQGSFHCGLPGWVLAPN-GVSCVTMGFVSLGPPSPGPPDEEDKGEKGSSTVPRAA 498  
QY 478 PPSPTPGSTLTTPA 491  
Db 499 TASPTRGPEGTPKA 512

## RESULT 14

US-10-408-765A-1422  
; Sequence 1422, Application US/10408765A  
; Publication No. US20040101874A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660088.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1422  
; LENGTH: 652  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-1422

Query Match 20.4%; Score 578; DB 16; Length 652;  
Best Local Similarity 31.6%; Pred. No. 5e-33;  
Matches 175; Conservative 52; Mismatches 201; Indels 126; Gaps 31;

QY 6 QPGGSG-----CWEHDFALYPGATFLNASQICDGLRGLMTVRSSVAAD-----VI 53  
Db 17 QPGAGTGADTEAVVCGTACTYTAHSGKLSAAEAQNHCNQGNGNLATVKSKEEAHQVRVL 76  
QY 54 SLLNGDGGVGR--RLWIGLQLP--CGDPKRLGLRFGQWVTGNNTSYSRWALDLN 110  
Db 77 AQLRREALTARMKFWIGLQREKGLDPSL--PLKGFVWVGGEPTPYSNWHKELRN 134  
QY 111 GAPLCGPLCVA--VSAAEATVPSE--PIWEEQOC-----EVKADGFLCFEHPATCRPLA 161  
Db 135 SC--ISKRCVSLLLDLSQPLLPRLPKWSEGPCGSPGSGNIEGFVCFKFGMCRPLA 192  
QY 162 V-EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMC--TAPPGAQVQHW-- 217  
Db 193 LGPG-----QVYTTPTFTTSSLEAVPFAAANVA-----CGEGDKDETQSHYFL 239  
QY 218 -AREAPGAWD-----CSVENGGCEHAC--NAIPGAPRCQCPAGALQADGRS 261  
Db 240 CKEAPADVDMWSSGPLCVSPKYGKFNNGGCHQDCFEFGDGSFLCGCRPFGLLDLVT 299  
QY 262 CTASATQSCNDLCE--HFCVNPDPQSGYSKCMETGYRLAADQHRCEDVDDCILEPSPCP 319  
Db 300 C-ASRNPSSSPCRGGATCVLGP-HGKNYTCRCPOGYQLDSSQLDCVDVDEC--QDSPCA 355  
QY 320 QRCVNTQGGFECHCYNDYLDVGECEVPVDFPCFRANCEYQCOPLNQTSYLVCVCAEGFAP 379  
Db 356 QECVNTPGGFRCECWGY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395  
QY 380 PHEPHRCQMFNCQACPADCDPNTQAS--CECEGYIL--DDGFICTDIDEC--ENGGFC 433  
Db 396 -----CAQGC--TNTDGSFHCSEEGYVLAGEDGTQCQDVDCVGGGGLC 439  
QY 434 SGVCHNLPGTFECICGPDGALVRHIGTDCDSGKV-----DGGDSGSGE-----P 477



Db 440 DSLCFNTQGSFHGCLPGWLAPN-GVSCWTMGVSLGPPSGPPDEEDKGEKGSTVPRAA 498  
QY 478 PPSPTGSGTILTPPA 491  
Db 499 TASPTRGPGCTPKA 512

Job time : 43 secs

RESULT 15  
US-09-789-919-96  
; Sequence 96, Application US/09789919  
; Patent No. US20020064855A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor  
; APPLICANT: Moore, Kateri  
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM  
; TITLE OF INVENTION: CELLS AND USES THEREOF  
; FILE REFERENCE: 2275-1-005  
; CURRENT APPLICATION NUMBER: US/09/789,919  
; CURRENT FILING DATE: 2001-02-21  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 96  
; LENGTH: 652  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-789-919-96

Query Match 20.4%; Score 576; DB 9; Length 652;  
Best Local Similarity 31.4%; Pred. No. 7e-33;  
Matches 174; Conservative 53; Mismatches 201; Indels 126; Gaps 31;  
QY 6 QPGSQ-----CUEHDCFALYPGPATFLNASQICDGLRHLMTVRSSVAAD-----VI 53  
Db 17 QPGAGTGADTEAVVCGTACTAHSKLSAAEAQNHCMQNGNLA TVKSEEAQHVRVL 76  
QY 54 SLILNGDGGVGR--RLWIGLQLPPG-CGDPKRLGPLRGFWTGTGDNNTSYSRWRLDLN 110  
Db 77 AQLRREAALTARMSKFWIGLQREKGLDPSL--PLKGFVWYGGEDTYSNWHKELRN 134  
QY 111 GABLCGCLCVA--VSAEATVPSE-PIWEEQQC-----EVKADGFLCEFFHFPATCRPLA 161  
Db 135 SC--ISKRCVSLLLDLSQLLPNLPKWSGPGSGPSGPNIEGFVCKFSFKGMRPLA 192  
QY 162 V-EFGAAAAVSTYGTFFAARGADFOALPVGSSAAVAPLQLMC-TAPPGAVQGHW-- 217  
Db 193 LGPGP-----QVTTYTPFTTSSLEAVPFASAAVNA-----CGEGKDETQSHYFL 239  
QY 218 -AREAPGAWD-----CSVENGGCEHAC-NAIPGAPRCOCPAGALOADGRS 261  
Db 240 CKEKAPDVFWSGSGPLCVSPKYGCNFNNGGCHQDCFEFGDGSFLCGCRFRLDLVT 299  
QY 262 CTASATQSCNDLCE--HFCVNPDPQGSYSCMCETGYRLAADQHRCEDVDDCILEPSPCP 319  
Db 300 C-ASRNPCSSSPCKGGATCVLGP-HGKNYTCRCPPQGVQLDSSQLDCVDVDEC--QDSPCA 355  
QY 320 QRCVNTGGPECHCYPNYDLVDGCEVDPVDFRANCEYOCQFLNQTSLYLCVCAEGFAP 379  
Db 356 QECVNTPGGRCEWVGY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395  
QY 380 PHEPHRCQFCNQTACPADCPNTQAS--CECPGXYL--DDGFICTDIDEC--ENGGFC 433  
Db 396 -----CAQGC-TWIDGSHCSCEEGVLAGEDGTQCCQDVDECVPGGGLC 439  
QY 434 SGVCHNLPGTFECICGPDLSALVRHIGTDCDSKV-----DGGDSGSGE-----P 477  
Db 440 DSLCFNTQGSFHGCLPGWLAPN-GVSCWTMGVSLGPPSGPPDEEDKGEKGSTVPRAA 498  
QY 478 PPSPTGSGTILTPPA 491  
Db 499 TASPTRGPGCTPKA 512

Search completed: June 9, 2004, 08:59:53

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:44:08 ; Search time 16.5 Seconds  
(without alignments)  
2903.235 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516

Perfect score: 2830

Sequence: 1 APAEPQGGQCVEHDCFAL.....PSPTFGSTLTTPPAVGLVHSG 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: PIR 78.\*

1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description        |
|------------|--------|-------------|--------|-------|--------------------|
| 1          | 2830   | 100.0       | 575    | 1     | THHUB              |
| 2          | 1854   | 65.5        | 577    | 2     | thrombomodulin pre |
| 3          | 1092.5 | 38.6        | 356    | 2     | thrombomodulin pre |
| 4          | 416    | 14.7        | 1620   | 2     | thrombomodulin - b |
| 5          | 414    | 14.6        | 1574   | 2     | hypothetical prote |
| 6          | 373    | 13.2        | 2307   | 2     | MEGF6 protein - ra |
| 7          | 367.5  | 13.0        | 1184   | 2     | fibrillin-2 precu  |
| 8          | 367    | 13.0        | 2871   | 2     | fibrillin-2 precu  |
| 9          | 365    | 12.9        | 2918   | 2     | fibrillin-1 - bovi |
| 10         | 363    | 12.8        | 3002   | 2     | fibrillin-2 precu  |
| 11         | 361    | 12.8        | 2871   | 2     | fibrillin-1 precu  |
| 12         | 359    | 12.7        | 741    | 2     | fibrillin-1 precu  |
| 13         | 357    | 12.6        | 1221   | 2     | hypothetical prote |
| 14         | 331    | 11.7        | 589    | 2     | fibulin-2 precu    |
| 15         | 329.5  | 11.6        | 1964   | 2     | fibulin-1D precu   |
| 16         | 327.5  | 11.6        | 1712   | 2     | notch4 - mouse     |
| 17         | 324.5  | 11.5        | 689    | 2     | masking protein pr |
| 18         | 324.5  | 11.5        | 712    | 2     | fibulin, splice fo |
| 19         | 320.5  | 11.3        | 2321   | 2     | fibulin 1, splice  |
| 20         | 320    | 11.3        | 1394   | 2     | notch3 protein - h |
| 21         | 314.5  | 11.1        | 3507   | 2     | transforming growt |
| 22         | 313    | 11.1        | 601    | 2     | hypothetical prote |
| 23         | 313    | 11.1        | 683    | 2     | fibulin 1 precu    |
| 24         | 313    | 11.1        | 798    | 2     | fibulin 1 precu    |
| 25         | 312.5  | 11.0        | 1820   | 2     | hypothetical prote |
| 26         | 308    | 10.9        | 685    | 2     | latent transformin |
| 27         | 308    | 10.9        | 705    | 2     | fibulin, splice fo |
| 28         | 303.5  | 10.7        | 1251   | 2     | latent transformin |
| 29         | 296    | 10.5        | 2531   | 2     | notch-1 protein -  |

cell-fate determin  
Notch homolog prot  
transmembrane prot  
notch protein - fir  
notch 3 protein -  
extracellular prote  
hypothetical prote  
epidermal growth f  
epidermal growth f  
notch protein homo  
notch protein homo  
epidermal growth f  
Notch B protein -  
notch homolog - se  
Xotch protein - Af  
growth potentiati  
epidermal growth f

30 286 10.1 2471 2 A49128  
31 285 10.1 2352 2 T30201  
32 283.5 10.0 2437 2 S42612  
33 280.5 9.9 2703 1 A24420  
34 280 9.9 2318 2 S45306  
35 278.5 9.8 387 2 I38449  
36 277.5 9.8 511 2 T17298  
37 276.5 9.8 493 2 JC5621  
38 275 9.7 2531 2 S18188  
39 272 9.6 2555 2 A40043  
40 269.5 9.5 1217 1 EGM5MG  
41 263 9.3 1203 2 A49175  
42 259.5 9.2 2531 2 T31070  
43 256.5 9.1 2524 2 A35844  
44 254.5 9.0 674 2 I55476  
45 249.5 8.8 1133 1 EGRT

#### ALIGNMENTS

##### RESULT 1

##### THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence revision 12-May-1995 #text change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S39554

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyar

J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed acti

A:Reference number: A41442; MUID:88227301; PMID:2836377

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DBJ:D00210; NID:G220126; PIDN:BAA00149.1; PID:G220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the cD

A:Reference number: A28307; MUID:87317665; PMID:2819876

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:G339658; PIDN:AAA61175.1; PID:G339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kawah

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endot

A:Reference number: A29680; MUID:88004395; PMID:2820710

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SUZ>

A:Cross-references: GB:X05495; NID:G37123; PIDN:CAA29045.1; PID:G736251

A:Experimental source: lung endothelium

A:Note: part of this sequence, including the amino end of the mature protein, were dete

R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of th

A:Reference number: A27073; MUID:88024950; PMID:2822087

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <WEN>

A:Cross-references: GB:M16552; NID:G339656; PIDN:AAB59508.1; PID:G339657

A:Experimental source: placenta

A:Note: parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792; PMID:8390446

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.

Biochem. J. 295, 131-140, 1993  
 A;Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.  
 A;Reference number: S38954; MUID:94029900; PMID:8216207  
 A;Accession: S38954  
 A;Molecule type: protein  
 A;Residues: 475-491, 'X', 493-494 <GER>  
 A;Note: the residue designated 'X' was determined to be a Ser with covalently bound chon  
 R;Meininger, D.P.; Komives, E.A.  
 submitted to the Brookhaven Protein Data Bank, September 1995  
 A;Reference number: A67369; PDB:1ZQA  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R;Tullinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, August 1994  
 A;Reference number: A52804; PDB:1HIT  
 A;Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
 R;Hrabal, R.; Komives, E.A.; Ni, F.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A;Reference number: A65583; PDB:1FGD  
 A;Contents: annotation; conformation by (1)H-NMR, residues 427-444  
 R;Hrabal, R.; Komives, E.A.; Ni, F.  
 Protein Sci. 5, 195-203, 1996  
 A;Title: Structural resiliency of an EGF-like subdomain bound to its target protein, thr  
 A;Reference number: A58595; MUID:96276211; PMID:8745396  
 A;Contents: annotation; conformation by (1)H-NMR  
 C;Genetics:  
 A;Gene: GDB:THBD  
 A;Cross-references: GDB:119613; OMIM:188040  
 A;Map position: 20p11.2-20p11.2  
 A;Introns: #status absent  
 C;Complex: homodimer, urinary form  
 C;Function:  
 A;Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activat  
 A;Pathway: blood coagulation moderation  
 A;Note: the membrane-bound form is located on the endothelium luminal surface of arterie  
 C;Note: thrombin complexed with the membrane-bound form is subject to endocytosis  
 C;Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
 C;Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coag  
 e protein  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
 F;19-513/Domain: extracellular #status predicted <EXT>  
 F;19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
 F;24-167/Domain: C-type lectin homology <LCH>  
 F;177-199/Region: PEST sequence  
 F;201-233/Region: PEST sequence  
 F;245-280/Domain: EGF homology <EG1>  
 F;288-323/Domain: EGF homology <EG2>  
 F;329-362/Domain: EGF homology <EG3>  
 F;369-404/Domain: EGF homology <EG4>  
 F;408-439/Domain: EGF homology <EG5>  
 F;445-480/Domain: EGF homology <EG6>  
 F;485-513/Region: PEST sequence  
 F;517-539/Domain: transmembrane #status predicted <TMN>  
 F;540-575/Domain: intracellular #status predicted <INT>  
 F;47-115, 116, 382, 409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;174, 225, 411, 504/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F;245-256, 252-267, 280, 288-296, 292-308, 310-323, 329-340, 336-349, 351-362, 369-378, 374-38  
 F;334, 498/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F;342/Modified site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
 F;490, 492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 100.0%; Score 2830; DB 1; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 9e-165;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APAEPPGSGQVEHDCFEALYPCFATFLNASQICDGLRGLMTVRSVAADVISLLNGD 60  
 Db 19 APAEPPGSGQVEHDCFEALYPCFATFLNASQICDGLRGLMTVRSVAADVISLLNGD 78  
 QY 61 GGVGRRLWIGLQPLPGCGDPKRLGLRGLFQWVTGDNNTSYSRWARLDINGAPLCGLCV 120  
 Db 79 GGVGRRLWIGLQPLPGCGDPKRLGLRGLFQWVTGDNNTSYSRWARLDINGAPLCGLCV 138

QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAVEPGAAAAVSYTYGTPFA 180  
 Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHPATCRPLAVEPGAAAAVSYTYGTPFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAPGAWDCSVENGCGEHCACNA 240  
 Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWAREAPGAWDCSVENGCGEHCACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGSGYSQWCETGYRLAA 300  
 Db 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGSGYSQWCETGYRLAA 318  
 QY 301 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGCEVFPVPCFRANCEYQC 360  
 Db 319 DQHRCEVDVDCILLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGCEVFPVPCFRANCEYQC 378  
 QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTCACADCDPNTQASCECPGVLDDGFI 420  
 Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTCACADCDPNTQASCECPGVLDDGFI 438  
 QY 421 CTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVDGSDSGSGEPSPPS 480  
 Db 439 CTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDCDSKVDGSDSGSGEPSPPS 498  
 QY 481 PTFGSTLTTPPAVGLVHSG 498  
 Db 499 PTFGSTLTTPPAVGLVHSG 516

RESULT 2  
 A60501  
 thrombomodulin precursor - mouse  
 N;Alternate names: fetomodulin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text\_change 16-Jul-1999  
 A;Accession: S08488; A32001; A60501  
 R;Dittman, W.A.; Majerus, P.W.  
 Nucleic Acids Res. 17, 802, 1989  
 A;Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted mou  
 A;Reference number: S08488; MUID:89128454; PMID:2536925  
 A;Accession: S08488  
 A;Molecule type: mRNA  
 A;Residues: 1-577 <DIT>  
 A;Cross-references: EMBL:X14432; NID:gs4781; PID:CAA32597.1; PID:gs4782  
 R;Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
 J. Biol. Chem. 263, 15915-15922, 1988  
 A;Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetate s  
 A;Reference number: A32001; MUID:89008498; PMID:2844823  
 A;Accession: A32001  
 A;Molecule type: mRNA  
 A;Residues: 97-577 <D12>  
 A;Cross-references: GB:J04060  
 R;Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.  
 Dev. Biol. 140, 113-122, 1990  
 A;Title: Identification of fetomodulin, a surface marker protein of fetal development, a  
 A;Reference number: A60501; MUID:90292331; PMID:2162790  
 A;Accession: A60501  
 A;Molecule type: protein  
 A;Residues: 19-22;330-343;479-489;545-555;562-575 <IMA>  
 C;Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofactor  
 C;Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
 C;Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
 F;24-165/Domain: C-type lectin homology <LCH>  
 F;244-279/Domain: EGF homology <EG1>  
 F;287-322/Domain: EGF homology <EG2>  
 F;328-361/Domain: EGF homology <EG3>  
 F;368-403/Domain: EGF homology <EG4>  
 F;407-438/Domain: EGF homology <EG5>  
 F;444-479/Domain: EGF homology <EG6>

Query Match 65.5%; Score 1854; DB 2; Length 577;  
 Best Local Similarity 65.8%; Pred. No. 1.6e-105;

## RESULT 4

submitted to the EMBL Data Library. September 1999

A;Reference number: Z20336  
A:Accession: T27303

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Residues: 1-1620 <WIL>

A: Experimental source: c

C;Genetics:  
A:Geno: GENO.VC40107 F

A; Introns: 77/1; 116/1;

Query Match

Matches 128: Conserve

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

100

22-10000

DB I16 D-ANEC--MAL

QY 212 AVQGHWAREA

Db 146 FELSGDNTC:

Qy 271 --NDLCEHFC"

7b 204 TNNCCCCC

100

Figure 1. Schematic diagram of the experimental setup. The subject is seated in a chair, viewing a screen displaying a target (a red dot) and a starting point (a green dot). The subject's hand is positioned at the starting point, and the target is located at a distance of 10 cm from the starting point. The subject is instructed to move the hand to the target as quickly and accurately as possible. The screen is positioned at a distance of 10 cm from the starting point. The subject's hand is positioned at the starting point, and the target is located at a distance of 10 cm from the starting point. The subject is instructed to move the hand to the target as quickly and accurately as possible. The screen is positioned at a distance of 10 cm from the starting point.

[illegible]

QY / QMF --- CNY

Db 321 EEMGGCQVGI

Qy 416 -----

Db 381 SYETVTCVTP"

Q. 430 UNIT DEFENDANT

7

100

## RESULT 5

TI3954  
MECE 304015

C;Species: Rattus norvegicus (Norway rat)  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: Tl3954  
R;Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A;Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A;Reference number: Z14126; MUID:98360089; PMID:9693030  
A;Accession: Tl3954  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1574 <NA>  
A;Cross-references: EMBL:AB011532; NID:g3449293; PID:BA032462.1; PID:G3449294  
A;Experimental source: strain Sprague-Dawley; brain  
C;Genetics:  
A;Gene: MEGF6

Query Match 14.6%; Score 414; DB 2; Length 1574;  
Best Local Similarity 34.1%; Pred. No. 9.7e-18;  
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

[illegible]

RESULT 6  
A57278  
fibrillin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibrillin genes suggests heterogeneity of extracellular matrix components in the developing mouse embryo.  
A:Reference number: A57278; PMID:95363670; PMID:7744963

Accession: A57278  
Status: preliminary  
Molecule type: DNA  
Residues: 1-2907 <ZHA>  
Cross-references: GB:U39790; NID:G762830; PID:AAA74908.1; PID:G762831  
Superfamily: fibrillin; EGF homology  
1239-1274/Domain: EGF homology <EGF>  
2488-2523/Domain: EGF homology <EGF>

Query Match 13.2%; Score 373; DB 2; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 5e-15;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

226 DCSVNGGCBHACNAIPGAPPCOPAGAAALQAGRSTASATSCNDLCEHFVCVPNPD-- 283  
          :  
1238 ECMIMNGGCDTQTNSRGSYECSCSEGVALLMPDRSCA-----DIDE--CENNPDIIC 1287

284 -----QPQSGYSCMCTGTGYRLAADHRCEDVDDCILEPSPCP-QRCVNTQGGECHCYCP 335

1288 DGGQCTNIPGEYRCLCYDGFMA5MDMKTCIDVNECDLNPNICMGEGENTKGSFICHQQL 1347  
QY 336 NYDLVDGE--CVEPVDPC--FRANCEYQOPLN-QTSYLCVCAEGFA-----PIP 380  
Db 1348 GYSVRKGTGTCTD-VDECEIGAHCNDMHASCLNVPGSFKSCREGWVGNGIKICDLDECA 1406  
QY 381 HEHRCOMFCNQCTACPADCPNTQAS--CECEGYILDDGFICTDDE-----CENG- 430  
Db 1407 NGTHQCSI-----NAQC-VNTPGSYRCACSEGF-TGDGFTCSDDVECAENTNLCENGQ 1457  
QY 431 -----GF-----CS-----GVCHNLPGETFCICGPD5ALV 455  
Db 1458 CLNVPGAYRCBCEMGFTPASDSRSCQIDECFSQNICVFSTCNNLPQMFHICIDGYYELD 1517  
QY 456 RHIG--ITCD 463  
Db 1518 RTGNCTDID 1527

RESULT 7  
A55184  
fibulin-2 precursor - human  
N:Alternate names: protein DKF2p586A1519.1  
C:Species: Homo sapiens (man)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 08-Sep-2002  
C:Accession: A55184; T08744  
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene  
A:Reference number: A55184; MUID:95104855; PMID:7806230  
A:Accession: A55184  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1184 <ZNA>  
A:Cross-references: GB:X82494; NID:9575232; PID:CAA57876.1; PID:G575233  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z16471  
A:Accession: T08744  
A:Molecule type: mRNA  
A:Residues: 636-719, QDECLMGANDCSRRQFCVNTLGSFYCVNHTVLCAQGYILNAHRKCVD', 720-853, 'T', 855-11:  
A:Cross-references: EMBL:AL050095  
A:Experimental source: adult uterus; clone DKF2p586A1519  
C:Genetics:  
A:Gene: GDB:FBLN2  
A:Cross-references: GDB:293037; OMIM:135821  
A:Map position: 3p25-3p24  
A:Note: DKF2p586A1519.1  
C:Superfamily: fibulin-2; EGF homology; von Willebrand factor type C repeat homology  
F:1-27/Domain: alternative splicing; extracellular matrix  
F:28-1184/Product: fibulin-2 protein #status predicted <SIG>  
F:905-941/Domain: EGF homology <EGF>

Query Match 13.0%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 5e-15;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

QY 136 EEQCEVKAD-GFLCEPHFATC-----RPLAV-----EPGAAAAVSITYGTTPAARG 183  
Db 531 EGQSCSNPLNGVPCN-HVMLSCEGEPLTVPEVRPEPFAAARRVS-----EAEH 582  
QY 184 ADPQALPVGSSAAV---APLGLQMLCTAPGAVQGHWAREAPGAWDCSVENGCGEACNA 240  
Db 583 AGREALSLGTEALPNSLPGDDQDECLLPDEL-----COHLCTN 622  
QY 241 IPGAPRCQCPAGALQADGRSC-----TASAT-----QS 269  
Db 623 TVGSYHCACFPFGSLQDDGRTCRPEGHPPQPEAPQEPALKSEFSQVANSNTIPLPLPQNT 582  
QY 270 CND--LCEHFCVNPDPQGSYCMCTGYRLAABQHRCEVDVDDCILEPSPCF--QRCVNT 325

Db 683 CKDNGPKQVC---STVGGSAICSPFGYAIMADGVSCDINECVTLDTCSRGEHVNT 739  
QY 326 QGQFECH---CYPNYDLVDGECVEPVDPFCFRANCEYQCCPLNQTSLVCV----- 371  
Db 740 LGSFHCYKALTCBPGYALKDGECE-EDVDEC--AMGTHTCQP-----GFLQNTKGSFYCOA 792  
QY 372 ---CARGFAPIPH-----EPHRCQMFNCOTACPADCDPNTQASCEPEGY-I 414  
Db 793 RORCMDGFLDPEGNCVDINECTSLSEPCRPFGSCINTVGSYTCQRPFLI---CARGYHA 849  
QY 415 LDGDFICTDDEGNGFCSCG---VCHNLPTGTETECIC 448  
Db 850 SDDGAKCVDVNECETVHRCGEGVCHNLPFGSYRCDC 886

RESULT 8  
A55567  
fibrillin 1 - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 24-Nov-2003  
C:Accession: A55567  
R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.  
Genomics 23, 480-485, 1994  
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to  
A:Reference number: A55567; MUID:95137597; PMID:7835900  
A:Accession: A55567  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <TIL>  
A:Cross-references: GB:L129748; NID:9508427; PIDN:AAA74122.1; PID:9508428  
C:Superfamily: fibrillin; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>

Query Match 13.0%; Score 367; DB 2; Length 2871;  
Best Local Similarity 31.5%; Pred. No. 1.2e-14;  
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 226 DCSVNGGGEHCNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCHVNP--- 282  
Db 1200 ECSIMNGGCTCTNTEGSEYSCQCPGFPALPDRSCT-----DIDE--CEDNPNIC 1249  
QY 283 -----DQPGSYSCMETGYRLAADHRCEDVDDCILEPSPC-PQRCVNTQGGFCHVNP 335  
Db 1250 DGGQCTNIPGEYRCLCYDGFMASEBDMKTCVDVNECDLNEINCLSGTCTKGSFICHCDM 1309  
QY 336 NYDLVDGE--CPEVPDPC--FRANCEYQCCPLNQT--SYLVCVCAEFGA-----PIP 380  
Db 1310 GYSKKGKGTGCTD-INECEIGAHCNDRHACVNTAGSFKSCSPGWIGIKCTDLDECS 1368  
QY 381 HEPHRCQMFNCOTACPADCDPNTQAS--CECEGYLDDGFTCTDIDE-EN----- 429  
Db 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDESENLNLCNGQ 1419  
QY 430 -----GGF-----CS-----GVCHNLPTGTETECICGPDALV 455  
Db 1420 CLNAGGGRCECDMGFVSADGKACEDIDECSLPNCVFGTCHNLPGLFRCEIGYELD 1479  
QY 456 RHIG-----TDCDSGK-VDGSDSGSGEPPPTPGSTLTPPAVLV 495  
Db 1480 RSGNCTDWECLDPTTCTISGNVNTPGSYTCDCPPD-----FELNPTRVGCV 1527

RESULT 9  
A54105  
fibrillin-2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 24-Nov-2003  
C:Accession: A54105; S17063; S31101  
R:Zhang, H.; Apfelloth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham,  
J. Cell Biol. 124, 855-863, 1994  
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe  
A:Reference number: A54105; MUID:94165150; PMID:8120105  
A:Accession: A54105

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr  
A:Molecule type: mRNA  
A:Residues: 1-2918 <ZHA>  
A:Cross-references: GB:U03272  
R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differ  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17063  
A:Molecule type: mRNA  
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
R:Miliewicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S31101  
A:Accession: S31101  
A:Molecule type: mRNA  
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 19  
A:Cross-references: EMBL:X62009  
C:Genetics: GDB:FBN2  
A:Gene: GDB:FBN2  
A:Cross-references: GDB:128122; OMIM:121050  
A:Map position: 5q23-5q31  
C:Superfamily: fibrillin; EGF homology  
C:Keywords: extracellular protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-2918/Product: fibrillin-2 #status predicted <MAT>  
F:1245-1280/Domain: EGF homology <EGF1>  
F:1970-2013/Domain: EGF homology <EGF>

Query Match 12.9%; Score 365; DB 2; Length 2918;  
Best Local Similarity 33.1%; Pred. No. 1.5e-14;  
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

QY 226 DCSVNGGGEHCNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCHVNP--- 283  
Db 1244 ECMINNGGCTCTNTEGSEYSCSEGVALMPDRSCA-----DIDE--CENPDTIC 1293  
QY 284 -----DQPGSYSCMETGYRLAADHRCEDVDDCILEPSPC-PQRCVNTQGGFCHVNP 335  
Db 1294 DGGQCTNIPGEYRCLCYDGFMASEBDMKTCVDVNECDLNEINCLSGTCTKGSFICHQOL 1353  
QY 336 NYDLVDGE--CPEVPDPC--FRANCEYQCCPLNQT--SYLVCVCAEFGAIPH-----EPH 385  
Db 1354 GYSVKKGTGCTD-VDECEIGAHCNDRHACVNTAGSFKSCREGW--IGNKIKCIDLDE 1410  
QY 386 COMFCNQTACPADCDPNTQAS--CECEGYLDDGFTCTDIDE-----CENG----- 430  
Db 1411 CSNGTHQCSINAQC-VNTPGSYRCACSEGF-TGDGFTCTDDECAENLNCNGQCLNVP 1468  
QY 431 -----GF-----CSGVCHNLPTGTETECICGPDALVHRHIG- 459  
Db 1469 GAYRCECEMGFTPASDRSCQDIDECSPONICVSGTCNNLFGMFHCICDDGYELDRTOGN 1528  
QY 460 -TDCC 463  
Db 1529 CTDID 1533

RESULT 10  
A47221  
fibrillin 1 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 24-Nov-2003  
C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198  
R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.  
Genomics 17, 476-484, 1993  
A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain struct  
A:Reference number: A47221; MUID:94010947; PMID:7691719  
A:Accession: A47221  
A:Molecule type: mRNA  
A:Residues: 1-337, 'T', 339-1029 <COR>  
A:Cross-references: GB:X63556

R;Pereira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangilinan, T.; Bonad  
Hum. Mol. Genet. 2, 961-968, 1993  
A;Title: Genomic organization of the sequence coding for fibrillin, the defective gene F  
A;Reference number: I54355; MUID:93372860; PMID:8364578  
A;Accession: I54355  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 132-3002 <PER>  
A;Cross-references: GB:113923; NID:G306745; PIDN:AAB02036.1; PID:G306746  
R;Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991  
A;Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A;Reference number: S17064; MUID:91304568; PMID:1852207  
A;Accession: S17064  
A;Molecule type: mRNA  
A;Residues: 1030-3002 <MAS>  
R;Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A;Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A;Reference number: I59574; MUID:93157831; PMID:8430317  
A;Accession: I59574  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 2217-2288, 'I', 2290-2325 <RES>  
A;Cross-references: GB:S54426; NID:G264860; PIDN:AAB25244.1; PID:G264861  
R;Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A;Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe  
A;Reference number: S17062; MUID:91304567; PMID:1852206  
A;Accession: S17062  
A;Molecule type: mRNA  
A;Residues: VLVTVVFIFLSYNKML, 944-1444 <LEE1>  
A;Cross-references: EMBL:X62008; NID:G31398; PIDN:CAB56534.1; PID:G5924015  
A;Accession: S62111  
A;Molecule type: protein  
A;Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LSE2>  
R;Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A;Title: Connective tissue microfibrils. Isolation and characterization of three large F  
A;Reference number: A34198; MUID:90078246; PMID:2512293  
A;Accession: A34198  
A;Molecule type: protein  
A;Residues: 565-575; 1890-1892, 'I', 1894-1900 <MAD>  
A;Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
C;Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C;Genetics: GDB:PBNI  
A;Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A;Map position: 15q21.1-15q21.1  
A;Introns: 2236/1; 2258/1; 2297/1  
C;Superfamily: fibrillin; EGF homology  
C;Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M  
F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted  
F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MAD>  
F:1332-1367/Domain: EGF homology <EGF>  
F:1457-1492/Domain: EGF homology <EGF2>  
F:2262-2295/Domain: EGF homology <EGF1>

QY 381 HEHRCOMFCNQACPADCDPNQAS--CECEGYILDDGFICTDDEC-EN----- 429  
Db 1500 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCTDLDECSENLCGNGQ 1550  
QY 430 -----GGF-----CS-----GVCHNLPGTECICGPPDSALV 455  
Db 1551 CLNAPGYRCECDMGVFPADGKACBIDECSLPNTCVFGTCHNLPLGFRCECEIYELD 1610  
QY 456 RHIG-----TDCDSGKVDGDSGSGEPSPPTPGS-----TLTPPAVGLV 495  
Db 1611 RSGGCTDVNECLDPTTCISGCVN-----TFGSYICDCPPFELNPTRVGCV 1658

RESULT 11  
A55624  
fibrillin-1 precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 24-Nov-2003  
C;Accession: A55624  
R;Yin, W.; Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Pereira, L.; Ramirez, I  
J. Biol. Chem. 270, 1798-1806, 1995  
A;Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin ge  
A;Reference number: A55624; MUID:95130561; PMID:7829516  
A;Accession: A55624  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-2871 <YIN>  
A;Cross-references: GB:L29454; NID:G575509; PIDN:AAA56840.1; PID:G575510  
C;Genetics:  
A;Gene: Fbn-1  
A;Superfamily: fibrillin; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.8%; Score 361; DB 2; Length 2871;  
Best Local Similarity 29.4%; Pred. No. 2.7e-14;  
Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;

QY 194 SAAVAPLGI-----QLMCTAPPAGVQGHAREAPGAWDCSVENGCGHACNAIFGAPRC 247  
Db 1162 SANLCFHGRCVNLIGKYQCACNPGYHPTHDLFCVDIDECSTMGCGCTCTCTNSDGYEC 1221  
QY 248 QCPAGAAALQADGRSCTASATQCNLCBHFVCPNP-----DOPGSYSCMCEGTGYRL 298  
Db 1222 SCQGFALMPQSRCT-----DIDQ--CEDNPNICDGGQCTNIPGEYRCLCYDGFMA 1271  
QY 299 AADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECHCYPNYDLVDGE--CVBPVDPDC--FR 353  
Db 1272 SEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDMGYSKKGKGTCTD-INECEIGA 1330  
QY 354 ANCEYOCQPLNOT-SYLCVCAEGFA-----PIPEHRCOMFCNQACPADCDPN 402  
Db 1331 HNGCRHVAVCTNTAGSFKCSGPGWIGDKICTDLDECSTMGTHMCSQH-----ADC-KN 1382  
QY 403 TQAS--CECEGYILDDGFICTDDEC-EN-----GGF----- 432  
Db 1383 TMGSYRCLCKDGY-TGDGFTCTDLDECSENLCGNGQCLNAPGYRCECDMGFVPSADG 1441  
QY 433 -----CS-----GVCHNLPGTECICGPPDSALVRHIG-----TDCDSGK 466  
Db 1442 KACEDIDECSLPNTCVFGTCHNLPLGFRCECEIGVELDRSGNCTDVNECLDPTTCISGN 1501  
QY 467 VDGDGSGSGEPSPPTPGS-----TLTPPAVGLV 495  
Db 1502 CVN-----TFGSYTCDCSPDFELNPTRVGCV 1527

RESULT 12  
T46488

hypothetical protein DKFZp434J065.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C;Accession: T46488

QY 226 DCSVENGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCBHFVCPNP--- 282  
Db 1331 ECSIMWGGCTCTNSSEGSYECSCQGFALMPDQSRCT-----DIDE--CEDNPNIC 1380  
QY 283 -----DOPGSYSCMCEGTGYRLAADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECHCYP 335  
Db 1381 DGGQCTNIPGEYRCLCYDGFMASEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDM 1440  
QY 336 NYDLVDGE--CVBPVDPDC--FRANCEYOCQPLNOT-SYLCVCAEGFA-----PIP 380  
Db 1441 GYSGKKGKGTCTD-INECEIGHNCGKHAVCTNTAGSFKCSGPGWIGDKIKCTDLDEC 1499



R;Duesterhoft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23035  
A:Accession: T46488  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-741 <AAA>  
A:Cross-references: EMBL:ALJ37638  
A:Experimental source: adult testis; clone DKFZp43J065  
C:Genetics:  
A:Note: DKFZp43J065.1

Query Match 12.7%; Score 359; DB 2; Length 741;  
Best Local Similarity 31.6%; Pred. No. 1.1e-14;  
Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;  
QY 227 CSVENGGCEHACNAIPGAPRCQCPAGALQDGRSCTA-SATQSCNDLCEHFVCPNPDQP 285  
DB 68 CAMEHDNCEQLCVNVPQSFVCCQYSGVALAEDGKRCVAVDYCASENHGCEHCY-NAD-- 124  
QY 286 GSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGPECHCYNYDL-VDGEC 344  
DB 125 GSYLCQCEHGFALNPEDEKCTKIDYCASSNHGCOHCEVNTDDSYSCHLKGFILNPKKT 184  
QY 345 VEPVDPCE--RANCEYQCPINQTSYLCVCAEGFAPIPH-----EPHRCQMFC 390  
DB 185 CRRINYCALNKPGEHECVNMEB-SYYCRCHGYTLDPNGKTCRVBDHCAQDQHGCEQLC 243  
QY 391 NQATCAPDCPNQAS--CECPGTYLDDGF-ICTDIDEC--ENGFCGVCCHNUPGTF 444  
DB 244 -----LNTEDSFVCCSEGLNEDLKTCRVYDCLSDHG--CYSYCVNMDRSF 291  
QY 445 ECTCGPSALVRHIGTDCDSKVDG---GDSG 473  
DB 292 ACQC-PEGHVLRSDGKTC-AKLDSCALGDHG 320

## RESULT 13

A:Accession: A49457  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 08-Sep-2002  
C:Accession: A49457; S74095  
R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.  
J. Cell Biol. 123, 1269-1277, 1993  
A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with  
A:Reference number: A49457; MUID:94064787; PMID:8245130  
A:Accession: A49457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1221 <PAN>  
A:Cross-references: GB:X75285; NID:9437046; PIDN:CAA53040.1; PID:9437047  
R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.  
Eur. J. Biochem. 240, 427-434, 1996  
A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases  
A:Reference number: S74094; MUID:96439073; PMID:8841408  
A:Accession: S74095  
A:Molecule type: protein  
A:Residues: 236-238, 'X', 240-247, 260-275, 336-344, 'L', 346-361, 405-426, 566-568, 'EM', 569-589  
C:Superfamily: fibulin-2; EGF homology; von Willebrand factor type C repeat homology  
C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer  
F:942-978/Domain: EGF homology <EGF>

Query Match 12.6%; Score 357; DB 2; Length 1221;  
Best Local Similarity 29.5%; Pred. No. 2.2e-14;  
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;  
QY 136 BEQQCEVKAD-CFLCEFFHPATC-----RPLAV-----EPGAAAAYS-----IT 174  
DB 521 EQQSCSEPNLGPNC-HVMLSCCEGEPFLIVPEVRPPPEAAPPVSEMEMASREALS 579  
QY 175 YGT-----PFAAGADFQ--ALP-----YGSSAAVAPLGLQLM-----CTAPPG 211

DB 580 LGTEAEPLNSLPGDDQDECLMLPGECLQHLCLINTVSYRCACPFGLQDGRTRCDRPG 639  
QY 212 AVQGHWARE-APGAWDCSV-----ENGGEHACNAIPGAPRCQCPAGAA 254  
DB 640 APQLDTARESA PRSESAQVSPNTIPLVPQNPCTKONGPCRCQVVGDTAMSCFFGYA 699  
QY 255 LQADGRSC-----TASATQSCNDLCEHFVCPNPDQPGSYSC-----MCETGYRLAADQHR 304  
DB 700 IMADGVSCEDQDECLMTHDCS--WKQFCV---NTLGSFYCVNHTVLCAGYILNA-HRK 753  
QY 305 CEDVDCCILEPSPC--PORCVNTQGGPECH---CYPNYDLVDEGCEVPEVDPCEFA--NC 356  
DB 754 CVDINECVTLHLCTRAEHCVNTGPGFQCYKALTCEBGVLTDECTD-VDECVTGFHNC 812  
QY 357 E--YQCPQLNQTSYLCV---CAEGFAPIPH-----EPHRCQMFCNQATCPA 397  
DB 813 QAGFSQCN-TKGSFYQARQRCMDGFLQDPGECVNDINECTSLLEPCRSFSCINTVGSY 871  
QY 398 DCDPNTQASCECEPEGY-ILDDGFICTDIDCEGNGFCSG---VCHNLPGTFECICGP 450  
DB 872 TCORNPLV---CGRGYHANESEGVVDVNBCEGTGVHRCGEGQLCVNLPFGSYRCDCKP 925

## RESULT 14

T43210  
fibulin-1D precursor - Caenorhabditis elegans (fragment)  
C:Species: Caenorhabditis elegans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 08-Sep-2002  
C:Accession: T43210  
R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.  
submitted to the EMBL Data Library, June 1998  
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character  
A:Reference number: Z22337  
A:Accession: T43210  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-589 <BAR>  
A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1  
C:Genetics:  
A:Note: Intron positions not resolved (incomplete sequence)  
C:Superfamily: fibulin-1; EGF homology

Query Match 11.7%; Score 331; DB 2; Length 589;  
Best Local Similarity 28.1%; Pred. No. 4.6e-13;  
Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

QY 85 GELRGF-QWVTGDNNTSYSRMARLDLNGAPLCPLCAVSAAEATVPSPFI-----WEQ 138  
DB 19 GCLRSFNKCCNGDIEITH---ASEIITGELPLNDPVLHLGDRCSHCHLCHDRGGEKV 75  
QY 139 QCEVKA-----DGFLCFHFPF--ATCRPLAVE-----PGAAAAAVSITYTTPFA--- 180  
DB 76 ECSCRSFGDLAPDGNACVDHIDECATLMDCCLESQRCCLNTPGSPKICRTLSGCTGYAMDS 135  
QY 181 --ARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQ-----GHWAREAPGAWDCS--- 228  
DB 136 ETERCEDVDECNLGH---DCGPLYQCRNTQGSYRCDAKKCGDELONPMTGETSITC 191  
QY 229 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGAAQADGRSCTA 264  
DB 192 PNGYYPKNGMNDIDECVTGHNCGAGECVNTGFSRCQCKGNLCAHYEVN----- 243  
QY 265 SATQSCNDL--CEH-----FCVNPDPQGSYSCMCETGYRLAADQHRCEVDVDCIL--- 313  
DB 244 GATGFCEDWNECQGVCGSMCEI---NLPGTYKCKGPGGYEFNDKAKKCEVDVDECIKAG 300  
QY 314 BPSFCPQRCVNTQGGFECHCYVNYDLV-DGCEVPEVDPCEFR--ANCEYQCPQLNQTSYLC 370  
DB 301 HVCDSLAEICNTIGSFCECKKPGQLASDGRRCEDVNECTTGIAACEQKCVNI-PSGYQC 359  
QY 371 VCAEGFAPIP-----HEPHRCQMFC--NQATCAPDCPNQAS--CECEPEGY-ILDDGFI 420  
DB 360 ICDRGFALGPDGDKCEDIDECISWAGSGNDLCMGCC--INTKGSYLCQCPGKYKIQPDGRT 418



QY 421 CTDIDECENGFCSCG---VCHNLPOTFEC 446  
Db 419 CVDVDECAMGE-CAGSDKVCVNTLGSFKC 446

RESULT 15  
T09059  
notch4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 08-Sep-2002  
C:Accession: T09059  
R:Rowen, L.; Mahairas, G.; Olin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; So  
submitted to the EMBL Data Library, October 1997  
A:Description: Sequence of the mouse major histocompatibility locus class III region.  
A:Reference number: Z16543  
A:Accession: T09059  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1964 <ROW>  
A:Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564947  
C:Genetics:  
A:Gene: notch4  
A:Map position: 17  
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67  
1679/3; 1729/1; 1761/3  
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
C:Keywords: receptor; signal transduction  
F:514-545/Domain: EGF homology <EGF>

Query Match 11.6%; Score 329.5; DB 2; Length 1964;  
Best Local Similarity 26.3%; Pred. No. 1.6e-12;  
Matches 128; Conservative 33; Mismatches 143; Indels 183; Gaps 31;

QY 116 GPLCVAVSAAEATVPSEPIWEBOQCEYKADGFLCE-FHFPATCRPLAVEPGAAAAVSIT 174  
Db 35 GGTCLRLSRGGIC-----QC---APGFLGETCQFPDPCR----- 66

QY 175 YGTFPFAARGADFOAL---PVGSSAAVAPLGLQLMCTAPPG---AVQGHWAREAPGAWDC 227  
Db 67 -DTQLCKNGGSCQALLTPPSSRSPTSLPHFSCFPCPGFTGDRQTHLEELCPFSF-C 124

QY 228 SVENGCGEACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQG- 286  
Db 125 S---NGG-HCYVQASGRPCQSCFPGWT---GEQCO-----LRDFCSANPCANG 166

QY 287 -----SYSCMETGYRLAADHRCF-DVDDCILEPSPCPQ--RCVNTQGGFECCHC--- 333  
Db 167 VCLATYPQIQCRPPGF-----EGHTCERDINECFLEFPQCPQTSCHNLTGSYQCLCPVG 222

QY 334 --YPNYDLVDGCEVPVDFCFRANCEYQCOPL---NQTSLVCARGFAPIPHE----- 382  
Db 223 QSGPQCKLRGAC--PPGSLNGG---TCQLVPEGHSTFHLCLCPGFTGLDCENMPDC 277

QY 383 -PHRCMFQNTACADCPDNTQASCECPGYLDDGFICT-DIDCE-----NGGF 432  
Db 278 VRHQCO---NGATCLDGLDTY---CLCPKW---KGWDCSEDIDECARGPPRCRNGGT 328

QY 433 C-----SGVCHNLPGTFCICGP----- 450  
Db 329 QNTAGSFHCVCVSGWGAGCEENLDDCAATCAPGTCIDRVGSFCLCPGRTGLLCH 388

QY 451 --DSALVR--HIGTDCDSKVDG-----GDSGS-----GEPPSP----- 481  
Db 389 LEDMCLSQPCHVNAQSTNPLTGLSTLCIQPGYSGSTCHQDLDECQMAQQGSPCEHGS 448

QY 482 ----TPGS 485  
Db 449 CINTPGS 455

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:42:38 ; Search time 10.5 Seconds  
(without alignments)  
2469.613 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516

Perfect score: 2830

Sequence: 1 APAEPQGGSCVHDFAL.....PSPTPGSLTPPAVLVHSG 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | ID           | Description         |
|------------|--------|-------------|--------|--------------|---------------------|
| 1          | 2826   | 99.9        | 575    | 1 TRBM_HUMAN | P07204 homo sapien  |
| 2          | 1854   | 65.5        | 577    | 1 TRBM_MOUSE | P15306 mus musculus |
| 3          | 1092.5 | 38.6        | 356    | 1 TRBM_BOVIN | P06579 bos taurus   |
| 4          | 578    | 20.4        | 652    | 1 CD93_HUMAN | Q9npv3 homo sapien  |
| 5          | 538    | 19.0        | 644    | 1 CD93_MOUSE | O89103 mus musculus |
| 6          | 521.5  | 18.4        | 643    | 1 CD93_RAT   | Q9et61 rattus norv  |
| 7          | 373    | 13.2        | 2907   | 1 FBN2_MOUSE | Q61555 mus musculus |
| 8          | 367.5  | 13.0        | 1184   | 1 FBL2_HUMAN | P98095 homo sapien  |
| 9          | 367    | 13.0        | 2871   | 1 FBN1_BOVIN | P98133 bos taurus   |
| 10         | 366    | 12.9        | 2871   | 1 FBN1_PIG   | Q9tv36 sus scrofa   |
| 11         | 365    | 12.9        | 2911   | 1 FBN2_HUMAN | P35556 homo sapien  |
| 12         | 363    | 12.8        | 2871   | 1 FBN1_HUMAN | P35555 homo sapien  |
| 13         | 361    | 12.8        | 2871   | 1 FBN1_MOUSE | O61554 mus musculus |
| 14         | 358    | 12.7        | 956    | 1 MTN2_HUMAN | O00339 homo sapien  |
| 15         | 357    | 12.6        | 1221   | 1 FBL2_MOUSE | P37889 mus musculus |
| 16         | 329.5  | 11.6        | 1864   | 1 NTC4_MOUSE | P31695 mus musculus |
| 17         | 328.5  | 11.6        | 956    | 1 MTN2_MOUSE | O08746 mus musculus |
| 18         | 328.5  | 11.6        | 1389   | 1 LTBS_MOUSE | O8cg18 mus musculus |
| 19         | 328.5  | 11.6        | 1713   | 1 LTBL_MOUSE | Q8cg19 mus musculus |
| 20         | 327.5  | 11.6        | 1712   | 1 LTBI_RAT   | Q00318 rattus norv  |
| 21         | 320.5  | 11.3        | 2321   | 1 NTC3_HUMAN | Q9um47 homo sapien  |
| 22         | 320    | 11.3        | 1394   | 1 LTBS_HUMAN | P22064 homo sapien  |
| 23         | 320    | 11.3        | 1595   | 1 LTBL_HUMAN | Q14766 homo sapien  |
| 24         | 315.5  | 11.1        | 443    | 1 FBL4_HUMAN | O95967 homo sapien  |
| 25         | 313    | 11.1        | 798    | 1 FBL1_HUMAN | P23142 homo sapien  |
| 26         | 313    | 11.1        | 798    | 1 FBL1_HUMAN | O77469 caenorhabdi  |
| 27         | 308.5  | 10.9        | 443    | 1 FBL4_CHICK | O55058 cricetus     |
| 28         | 308.5  | 10.9        | 2003   | 1 NTC4_HUMAN | Q99466 homo sapien  |
| 29         | 308    | 10.9        | 448    | 1 FBL5_HUMAN | Q9ubx5 homo sapien  |
| 30         | 307    | 10.8        | 704    | 1 FBL1_CHICK | Q73775 gallus gall  |
| 31         | 306.5  | 10.8        | 443    | 1 FBL4_MOUSE | Q9wvj9 mus musculus |
| 32         | 306    | 10.8        | 681    | 1 FBL1_BRARE | O42182 brachydanio  |
| 33         | 306    | 10.8        | 705    | 1 FBL1_MOUSE | Q08879 mus musculus |

## RESULT 1

| ID | TRBM_HUMAN   | STANDARD; | PRT; | 575 AA. |
|----|--|-----------|------|---------|
| AC | P07204;  |           |      |         |
| DT | 01-APR-1988 (Rel. 07, Created)   |           |      |         |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update)                            |           |      |         |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update)                          |           |      |         |
| DE | Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).           |           |      |         |
| GN | THBD OR THRM.  |           |      |         |
| OS | Homo sapiens (Human).  |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |           |      |         |
| OC | Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.               |           |      |         |
| OX | NCBI_TaxID=9606;   |           |      |         |
| RN | [1]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88004395; PubMed=2820710;                                      |           |      |         |
| RA | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I.,        |           |      |         |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;            |           |      |         |
| RT | "Structure and expression of human thrombomodulin, a thrombin          |           |      |         |
| RT | receptor on endothelium acting as a cofactor for protein C             |           |      |         |
| RT | activation."   |           |      |         |
| RL | EMBO J. 6:1891-1897(1987).   |           |      |         |
| RN | [2]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88024950; PubMed=2822087;                                      |           |      |         |
| RA | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.; |           |      |         |
| RT | "Human thrombomodulin: complete cDNA sequence and chromosome           |           |      |         |
| RT | localization of the gene."   |           |      |         |
| RL | Biochemistry 26:4350-4357(1987).                                       |           |      |         |
| RN | [3]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=87317665; PubMed=2819876;                                      |           |      |         |
| RA | Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;         |           |      |         |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences  |           |      |         |
| RT | of the cDNA and gene predict protein structure and suggest sites of    |           |      |         |
| RT | regulatory control."   |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).                      |           |      |         |
| RN | [4]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88227901; PubMed=2836377;                                      |           |      |         |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,              |           |      |         |
| RA | Deyashiki Y., Maruyama I., Suzuki K.;                                  |           |      |         |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin-      |           |      |         |
| RT | catalyzed activation of protein C."                                    |           |      |         |
| RL | J. Biochem. 103:281-285(1988).   |           |      |         |
| RN | [5]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RA | Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,       |           |      |         |
| RA | Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;                 |           |      |         |
| RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.                |           |      |         |
| RN | [6]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=21638749; PubMed=11780052;                                     |           |      |         |
| RA | Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,     |           |      |         |
| RA | Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,     |           |      |         |
| RA | Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,            |           |      |         |

## ALIGNMENTS

| ID | TRBM_HUMAN   | STANDARD; | PRT; | 575 AA. |
|----|--|-----------|------|---------|
| AC | P07204;  |           |      |         |
| DT | 01-APR-1988 (Rel. 07, Created)   |           |      |         |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update)                            |           |      |         |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update)                          |           |      |         |
| DE | Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).           |           |      |         |
| GN | THBD OR THRM.  |           |      |         |
| OS | Homo sapiens (Human).  |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |           |      |         |
| OC | Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.               |           |      |         |
| OX | NCBI_TaxID=9606;   |           |      |         |
| RN | [1]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88004395; PubMed=2820710;                                      |           |      |         |
| RA | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I.,        |           |      |         |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;            |           |      |         |
| RT | "Structure and expression of human thrombomodulin, a thrombin          |           |      |         |
| RT | receptor on endothelium acting as a cofactor for protein C             |           |      |         |
| RT | activation."   |           |      |         |
| RL | EMBO J. 6:1891-1897(1987).   |           |      |         |
| RN | [2]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88024950; PubMed=2822087;                                      |           |      |         |
| RA | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.; |           |      |         |
| RT | "Human thrombomodulin: complete cDNA sequence and chromosome           |           |      |         |
| RT | localization of the gene."   |           |      |         |
| RL | Biochemistry 26:4350-4357(1987).                                       |           |      |         |
| RN | [3]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=87317665; PubMed=2819876;                                      |           |      |         |
| RA | Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;         |           |      |         |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences  |           |      |         |
| RT | of the cDNA and gene predict protein structure and suggest sites of    |           |      |         |
| RT | regulatory control."   |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).                      |           |      |         |
| RN | [4]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88227901; PubMed=2836377;                                      |           |      |         |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,              |           |      |         |
| RA | Deyashiki Y., Maruyama I., Suzuki K.;                                  |           |      |         |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin-      |           |      |         |
| RT | catalyzed activation of protein C."                                    |           |      |         |
| RL | J. Biochem. 103:281-285(1988).   |           |      |         |
| RN | [5]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RA | Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,       |           |      |         |
| RA | Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;                 |           |      |         |
| RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.                |           |      |         |
| RN | [6]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=21638749; PubMed=11780052;                                     |           |      |         |
| RA | Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,     |           |      |         |
| RA | Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,     |           |      |         |
| RA | Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,            |           |      |         |

Q9wvh9 mus musculus  
Q9wvh8 rattus norv  
Q01705 mus musculus  
Q9r172 rattus norv  
Q35516 mus musculus  
Q8nj99 cercopithec  
Q12805 homo sapien  
Q9w30 rattus norv  
Q04721 homo sapien  
P46530 brachydanio  
P07207 drosophila  
Q61982 mus musculus

| ID | TRBM_HUMAN   | STANDARD; | PRT; | 575 AA. |
|----|--|-----------|------|---------|
| AC | P07204;  |           |      |         |
| DT | 01-APR-1988 (Rel. 07, Created)   |           |      |         |
| DT | 01-FEB-1991 (Rel. 17, Last sequence update)                            |           |      |         |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update)                          |           |      |         |
| DE | Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).           |           |      |         |
| GN | THBD OR THRM.  |           |      |         |
| OS | Homo sapiens (Human).  |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |           |      |         |
| OC | Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.               |           |      |         |
| OX | NCBI_TaxID=9606;   |           |      |         |
| RN | [1]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88004395; PubMed=2820710;                                      |           |      |         |
| RA | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I.,        |           |      |         |
| RA | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;            |           |      |         |
| RT | "Structure and expression of human thrombomodulin, a thrombin          |           |      |         |
| RT | receptor on endothelium acting as a cofactor for protein C             |           |      |         |
| RT | activation."   |           |      |         |
| RL | EMBO J. 6:1891-1897(1987).   |           |      |         |
| RN | [2]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88024950; PubMed=2822087;                                      |           |      |         |
| RA | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.; |           |      |         |
| RT | "Human thrombomodulin: complete cDNA sequence and chromosome           |           |      |         |
| RT | localization of the gene."   |           |      |         |
| RL | Biochemistry 26:4350-4357(1987).                                       |           |      |         |
| RN | [3]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=87317665; PubMed=2819876;                                      |           |      |         |
| RA | Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;         |           |      |         |
| RT | "Human thrombomodulin gene is intron depleted: nucleic acid sequences  |           |      |         |
| RT | of the cDNA and gene predict protein structure and suggest sites of    |           |      |         |
| RT | regulatory control."   |           |      |         |
| RL | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).                      |           |      |         |
| RN | [4]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=88227901; PubMed=2836377;                                      |           |      |         |
| RA | Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,              |           |      |         |
| RA | Deyashiki Y., Maruyama I., Suzuki K.;                                  |           |      |         |
| RT | "Gene structure of human thrombomodulin, a cofactor for thrombin-      |           |      |         |
| RT | catalyzed activation of protein C."                                    |           |      |         |
| RL | J. Biochem. 103:281-285(1988).   |           |      |         |
| RN | [5]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RA | Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,       |           |      |         |
| RA | Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;                 |           |      |         |
| RL | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.                |           |      |         |
| RN | [6]  |           |      |         |
| RP | SEQUENCE FROM N.A.   |           |      |         |
| RX | MEDLINE=21638749; PubMed=11780052;                                     |           |      |         |
| RA | Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,     |           |      |         |
| RA | Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,     |           |      |         |
| RA | Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,            |           |      |         |

RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clapp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekesch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levenshah M.H., Levenshah M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R.J., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suleston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RA "The DNA sequence and comparative analysis of human chromosome 20";  
RA Nature 414:865-871(2001).  
RA [7]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedini T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.S., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield F.S.N., Krzywinski M.I., Skaleka U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RA human and mouse cDNA sequences";  
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RA [8]  
RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RX MEDLINE=94029900; PubMed=8216207;  
RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,  
RA Grinnell B.W.;  
RA "Identification of the predominant glycosaminoglycan-attachment site  
RA in soluble recombinant human thrombomodulin: potential regulation of  
RA functionality by glycosyltransferase competition for serine474";  
RA Biochem. J. 295:131-140(1993).  
RA [9]  
RP STRUCTURE BY NMR OF 389-407.  
RX MEDLINE=96007474; PubMed=7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RA "The structure of a 19-residue fragment from the C-loop of the fourth  
RA epidermal growth factor-like domain of thrombomodulin";  
RA J. Biol. Chem. 270:23366-23372(1995).  
RA [10]  
RP STRUCTURE BY NMR OF 364-407.  
RX MEDLINE=96100636; PubMed=8529067;  
RA Meininger D.P., Hunter M.J., Komives E.A.;  
RA "Synthesis, activity, and preliminary structure of the fourth  
RA EGF-like domain of thrombomodulin";  
RA Protein Sci. 4:1683-1695(1995).  
RA [11]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE=95034791; PubMed=7947766;  
RA Sriniwasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RA "Thrombin-bound structure of an EGF subdomain from human  
RA thrombomodulin determined by transferred nuclear Overhauser  
RA effects";  
RA Biochemistry 33:13553-13560(1994).  
RA [12]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE=96276211; PubMed=8745396;  
RA Hrabal R., Komives E.A., Ni F.;  
RA "Structural resiliency of an EGF-like subdomain bound to its target  
RA protein, thrombin";  
RA Protein Sci. 5:195-203(1996).  
RA [13]  
RP STRUCTURE BY NMR OF 405-444.  
RX MEDLINE=98035729; PubMed=9367781;  
RA Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;  
RA "Structure of the fifth EGF-like domain of thrombomodulin: an  
RA EGF-like domain with a novel disulfide-bonding pattern";  
RA J. Mol. Biol. 273:913-926(1997).  
RA [14]  
RP VARIANT TED TYR-486.  
RX MEDLINE=95111115; PubMed=7811989;  
RA Oehlin A.-K., Marlar R.A.;  
RA "The first mutation identified in the thrombomodulin gene in a  
RA 45-year-old man presenting with thromboembolic disease";  
RA Blood 85:330-336(1995).  
RA [15]  
RP VARIANT TED TYR-486, AND VARIANTS THR-43; ALA-79; SER-495 AND LEU-501.  
RX MEDLINE=97341986; PubMed=9198186;  
RA Oehlin A.-K., Norlund L., Marlar R.A.;  
RA "Thrombomodulin gene variations and thromboembolic disease";  
RA Thromb. Haemost. 78:396-400(1997).  
RA [16]  
RP VARIANT VAL-473.  
RX MEDLINE=97206518; PubMed=9157575;  
RA Norlund L., Holm J., Zolner B., Oehlin A.-K.;  
RA "A common thrombomodulin amino acid dimorphism is associated with  
RA myocardial infarction";  
RA Thromb. Haemost. 77:248-251(1997).  
RA [17]  
RP VARIANT THR-43.  
RX MEDLINE=99057299; PubMed=9843165;  
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stubbs P.J., Manger Cats V., Ireland H.;  
RA "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
RA and the risk of myocardial infarction in men";  
RA Thromb. Haemost. 80:743-748(1998).  
RA [18]  
RP VARIANT VAL-473.  
RX MEDLINE=21143723; PubMed=11245641;  
RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RA "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
RA disease";  
RA Circulation 103:1386-1389(2001).  
RA [19]  
RP VARIANT TED TYR-486, AND VARIANT VAL-473.  
RX MEDLINE=22135346; PubMed=12139752;  
RA Faioni E.M., Franchi F., Castaman G., Biguzzi E., Rodeghiero F.;  
RA "Mutations in the thrombomodulin gene are rare in patients with  
RA severe thrombophilia";  
RA Br. J. Haematol. 118:595-599(2002).  
RA CC -!- FUNCTION: Thrombomodulin is a specific endothelial cell receptor  
RA that forms a 1:1 stoichiometric complex with thrombin. This  
RA complex is responsible for the conversion of protein C to the  
RA activated protein C (protein Ca). Once evolved, protein Ca  
RA scissions the activated cofactors of the coagulation mechanism,  
RA factor Va and factor VIIIa, and thereby reduces the amount of  
RA thrombin generated.  
RA CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
RA CC -!- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing  
RA thrombomodulin.

Query Match 99.9%; Score 2826; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 3.4e-180;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 60  
Db 19 APAEPQGGSCVHDCFALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 78  
Qy 61 GGVERRLWLTGLQPPCGGPKRLGRLGFWMTGDNNTSYSRWARLDLNGAPLCGLCV 120  
Db 79 GGVERRLWLTGLQPPCGGPKRLGRLGFWMTGDNNTSYSRWARLDLNGAPLCGLCV 138  
Qy 121 AVSAAETVSEPTWEEQCEVKADGFLCEFHFPATCRPLAVEPAAAAVSTYGTGPPA 180  
Db 139 AVSAAETVSEPTWEEQCEVKADGFLCEFHFPATCRPLAVEPAAAAVSTYGTGPPA 198  
Qy 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWRBAPGAWDCSVNGGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHWRBAPGAWDCSVNGGCEHACNA 258  
Qy 241 IPGAPRCQCPAGALQADGSCSTASATQSCNDLCEHFCVNPDPGSGYSCMCSTGYELAA 300  
Db 259 IPGAPRCQCPAGALQADGSCSTASATQSCNDLCEHFCVNPDPGSGYSCMCSTGYELAA 318  
Qy 301 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 378  
Qy 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCECEGVIDDGF 420  
Db 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPNTQASCECEGVIDDGF 438  
Qy 421 CTIDICENGFGSCGVCHNLPGTFECICGPDALVHIGTDCDSCGVKVDGSDSGSGRPPPS 480  
Db 439 CTIDICENGFGSCGVCHNLPGTFECICGPDALVHIGTDCDSCGVKVDGSDSGSGRPPPS 498  
Qy 481 PTPGSTLTTPPAVLVHSG 498  
Db 499 PTPGSTLTTPPAVLVHSG 516

RESULT 2

TRM\_MOUSE STANDARD; PRT; 577 AA.

AC F15306;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Thrombomodulin precursor (Fetomodulin) (TM).  
GN THBD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89128454; PubMed=2536925;  
RA Dittman W.A., Majerus P.W.;  
RT "Sequence of a cDNA for mouse thrombomodulin and comparison of the  
RT predicted mouse and human amino acid sequences.";  
RL Nucleic Acids Res. 17:802-802(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89008498; PubMed=2844823;  
RA Dittman W.A., Kumada T., Sadler J.E., Majerus P.W.;  
RT "The structure and function of mouse thrombomodulin. Phorbol  
RT myristate acetate stimulates degradation and synthesis of  
RT thrombomodulin without affecting mRNA levels in hemangioma cells.";  
RL J. Biol. Chem. 263:15815-15822(1988).  
CC -!- FUNCTION: Thrombomodulin is a specific endothelial cell receptor  
CC that forms a 1:1 stoichiometric complex with thrombin. This  
CC complex is responsible for the conversion of protein C to the

activated protein C (protein C). Once evolved, protein C  
scissions the activated cofactors of the coagulation mechanism,  
factor Va and factor VIIIa, and thereby reduces the amount of  
thrombin generated.  
-!- SUBCELLULAR LOCATION: Type I membrane protein.  
-!- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing  
thrombomodulin (By similarity).  
-!- SIMILARITY: Contains 6 EGF-like domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X14432; CAA32597.1; -;  
PIR; S08488; A60501.  
HSP; P07204; LEGT.  
MGD; MG198736; Thbd.  
GO; GO:000790; P:embryonic development; IMP.  
GO; GO:0007565; P:pregnancy; IMP.  
InterPro; IPR001152; Asx\_hydroxyl\_s.  
InterPro; IPR001881; EGF\_Ca  
InterPro; IPR006209; EGF\_Like.  
InterPro; IPR001304; Lectin\_C.  
InterPro; IPR001491; Thrombomodulin.  
Pfam; PF00008; EGF\_3.  
Pfam; PF00059; lectin\_c; 1.  
PRINTS; PR00907; THROMBOMODULN.  
SMART; SM00034; CLECT; 1.  
SMART; SM00179; EGF\_CA; 1.  
PROSITE; PS00010; ASX\_HYDROXYL; 2.  
PROSITE; PS0041; C\_TYPE\_LLECTIN\_2; 1.  
PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
PROSITE; PS01186; EGF\_2; 3.  
PROSITE; PS00026; EGF\_3; 3.  
PROSITE; PS01187; EGF\_CA; 2.  
XW Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;  
XW Glycoprotein; Signal; EGF-like domain.  
FT CHAIN 1 16  
FT SIGNAL 1 16  
FT POTENTIAL.  
FT THROMBOMODULIN.  
FT DOMAIN 17 517  
FT EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 518 541  
FT POTENTIAL.  
FT CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 542 577  
FT EGF-LIKE 1.  
FT DOMAIN 280 280  
FT EGF-LIKE 2.  
FT DOMAIN 324 323  
FT EGF-LIKE 3.  
FT DOMAIN 364 404  
FT EGF-LIKE 4.  
FT DOMAIN 403 439  
FT EGF-LIKE 5.  
FT DOMAIN 440 480  
FT EGF-LIKE 6.  
FT DISULFID 244 255  
FT BY SIMILARITY.  
FT DISULFID 251 264  
FT BY SIMILARITY.  
FT DISULFID 266 279  
FT BY SIMILARITY.  
FT DISULFID 287 295  
FT BY SIMILARITY.  
FT DISULFID 291 307  
FT BY SIMILARITY.  
FT DISULFID 309 322  
FT BY SIMILARITY.  
FT DISULFID 328 339  
FT BY SIMILARITY.  
FT DISULFID 335 348  
FT BY SIMILARITY.  
FT DISULFID 350 361  
FT BY SIMILARITY.  
FT DISULFID 368 377  
FT BY SIMILARITY.  
FT DISULFID 373 387  
FT BY SIMILARITY.  
FT DISULFID 389 403  
FT BY SIMILARITY.  
FT DISULFID 407 416  
FT BY SIMILARITY.  
FT DISULFID 412 424  
FT BY SIMILARITY.  
FT DISULFID 426 438  
FT BY SIMILARITY.  
FT DISULFID 444 454  
FT BY SIMILARITY.  
FT DISULFID 449 463  
FT BY SIMILARITY.  
FT DISULFID 465 479  
FT BY SIMILARITY.  
FT CARBOHYD 113 113  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 243 243  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 256 256  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).

|    |                           |              |                     |                                      |                             |
|----|---------------------------|--------------|---------------------|--------------------------------------|-----------------------------|
| FT | CARBOHYD                  | 408          | 408                 | N-LINKED (GLCNAC. . .) (POTENTIAL) . |                             |
| FT | CARBOHYD                  | 494          | 494                 | O-LINKED (GLYCOSAMINOGLYCAN) (BY     |                             |
| FT |                           |              |                     | SIMILARITY) .                        |                             |
| SQ | SEQUENCE                  | 577          | AA: 61867           | MW; B20E50B0FE745014 CRC64;          |                             |
|    | Query Match               |              | 65.5%               | ; Score 1854; DB 1; Length 577;      |                             |
|    | Best Local Similarity     |              | 65.8%               | ; Pred. No. 7.8e-116;                |                             |
|    | Matches 331; Conservative |              | 45; Mismatches 119; | Indels 8; Gaps 4;                    |                             |
| QY | 1                         | AAEPQPGSGSCV | HEHDCFALYPGPATF     | LNASQICDGLRGHLMTVR                   | SSVAADVISILLNGD 60          |
| DB | 19                        | ALAKLQPTGSGC | VEHECFALFQGPATF     | LDASOACRLQGHLMTVR                    | SSVAADVISLLLSQS 78          |
| QY | 61                        | G-GVGRRLWLGL | LPCCGDPKRLGPLRG     | QWTVGDNNTYSWARI                      | DLINGAPLCGPLC 119           |
| DB | 79                        | SMDLGP---    | WLGLQLPQCDDPV       | VHLGPIRGQWTVGDN                      | NHTSYSWARPNDQTAPLCGPLC 135  |
| QY | 120                       | VAVSAAENATPS | EPIWEFQECVCKADG     | FLCFEHFPATCRPLA                      | VEP-GAAAAVAVSIITGTP 178     |
| DB | 136                       | VTVSTATAAPGE | PAWEKPCETETQGL      | CFEYFTASCRPLTV                       | NTRPEAAHISSTWTP 195         |
| QY | 179                       | FAARGADFOALP | VGSSAAVAPLGLQ       | LCTAPPAGVQGHWA                       | REAPGAWDCSVENGCCSHAC 238    |
| DB | 196                       | FGVSGADFQTL  | PVSSAAVEPLGUEL      | VLCRAPPGTSEGHWA                      | EATGAWNCVSEVGGCGEYLC 255    |
| QY | 239                       | NAIFGAPRCQCP | AGAAALQADGRSCT      | ASATQSCNDLCEHF                       | CVNPDPQGSYSCMCETGYRL 298    |
| DB | 256                       | NRSTNEPRCLC  | PRDMDLQADGRS        | CARPVWQSCNELCE                       | HFCSVAEIVGYSYSCMCETGYQL 315 |
| QY | 299                       | AADQHRCEDVDD | CCILEPSPCORCV       | NTQGGFECHCYNY                        | LDVGECEVPEVDPFRANCEY 358    |
| DB | 316                       | AADGHRCEDVDD | CKQGNPCQLCVN        | TGKGFEPCYDGYE                        | LVDPGECEVELLDPCFGSNCEP 375  |
| QY | 359                       | QCOPLNOSTYLC | VCAEGFAPIPHEP       | HRKCMFCNQTCAP                        | CDPCDNTQASCEPCGYITLDDG 418  |
| DB | 376                       | QCQVPSPTDYRC | IACAPGAPKPDPE       | PHKCFMFCNETSC                        | FPADCDPNSPTVCECPGFIIDEG 435 |
| QY | 419                       | FICTIDECENG  | GFCSGVCHNLCT        | FPFCIPGPSALVR                        | HITGTCDSGKV----DGGDSGSG 475 |
| DB | 436                       | SVCTIDECQGC  | FTSECRNFGSYE        | ICICGPDOTALAQ                        | IISKDCDPIPVREDTKEEGSG 495   |
| QY | 476                       | EPPEPSPFGS   | TLTTPPAVLVHSG       | 498                                  |                             |
| DB | 496                       | EPFVSPTPG    | SFTGTPPPSARPVHSG    | 518                                  |                             |

### RESULT 3

|            |  |                                   |         |
|------------|--|-----------------------------------|---------|
| TRFM_BOVIN | STANDARD;  | PR7;                              | 356 AA. |
| ID         | AC   | P06579;                           |         |
| DT         | 01-JAN-1988  | (Rel. 06, Created)                |         |
| DT         | 01-JAN-1988  | (Rel. 06, Last sequence update)   |         |
| DT         | 16-OCT-2001  | (Rel. 40, Last annotation update) |         |
| DE         | Thrombomodulin (Fetomodulin) (TM) (Fragment).<br>THBD.               |                                   |         |
| GN         | Bos taurus (Bovine).   |                                   |         |
| OC         | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;              |                                   |         |
| OC         | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;    |                                   |         |
| OC         | Bovidae; Bovinae; Bos.   |                                   |         |
| OX         | NCBI_TaxID=9913;   |                                   |         |
| RN         | [1]  |                                   |         |
| RP         | SEQUENCE FROM N.A.   |                                   |         |
| RX         | MEDLINE=87067408; PubMed=3024152;                                    |                                   |         |
| RA         | Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;            |                                   |         |
| RT         | "Characterization of a thrombomodulin cDNA reveals structural        |                                   |         |
| RL         | similarity to the low density lipoprotein receptor.;"                |                                   |         |
| RL         | Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).                    |                                   |         |
| CC         | -1- FUNCTION: Thrombomodulin is a specific endothelial cell receptor |                                   |         |
| CC         | that forms a 1:1 stoichiometric complex with thrombin. This          |                                   |         |
| CC         | complex is responsible for the conversion of protein C to the        |                                   |         |
| CC         | activated protein C (protein Ca). Once evolved, protein Ca           |                                   |         |
| CC         | scissions the activated cofactors of the coagulation mechanism,      |                                   |         |
| CC         | factor Va and factor VIIIa, and thereby reduces the amount of        |                                   |         |

thrombin generated.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing

CC thrombomodulin.

CC -!- SIMILARITY: Contains 6 EGF-like domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M14657; AAA30785.1; --

DR PIR; A25918; A25918.

DR HSP; P07204; 1TWR.

DR InterPro; IPR000152; Asx\_hydroxyl\_S.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR001491; Thrombomodulin.

DR Pfam; PF00008; EGF; 4.

DR PRINTS; PR00907; THROMBOMODULN.

DR SMART; SM00179; EGF\_CA\_1.

DR PROSITE; PS00010; ASX\_HYDROXYL; 2.

DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.

DR PROSITE; PS01186; EGF\_2; 3.

DR PROSITE; PS00026; EGF\_3; 3.

DR PROSITE; PS01187; EGF\_CA; 2.

DR Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;

KW Glycoprotein; EGF-like domain.

FT NON\_TER 1 1

FT DOMAIN 1 296 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 297 320 POTENTIAL.

FT DOMAIN 321 356 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 17 57 EGF-LIKE 1.

FT DOMAIN 60 98 EGF-LIKE 2.

FT DOMAIN 99 137 EGF-LIKE 3.

FT DOMAIN 139 179 EGF-LIKE 4.

FT DOMAIN 178 214 EGF-LIKE 5.

FT DOMAIN 215 254 EGF-LIKE 6.

FT DISULFID 21 32 BY SIMILARITY.

FT DISULFID 28 41 BY SIMILARITY.

FT DISULFID 43 56 BY SIMILARITY.

FT DISULFID 64 72 BY SIMILARITY.

FT DISULFID 68 82 BY SIMILARITY.

FT DISULFID 84 97 BY SIMILARITY.

FT DISULFID 103 114 BY SIMILARITY.

FT DISULFID 110 123 BY SIMILARITY.

FT DISULFID 125 136 BY SIMILARITY.

FT DISULFID 143 152 BY SIMILARITY.

FT DISULFID 148 162 BY SIMILARITY.

FT DISULFID 164 178 BY SIMILARITY.

FT DISULFID 182 191 BY SIMILARITY.

FT DISULFID 187 199 BY SIMILARITY.

FT DISULFID 201 213 BY SIMILARITY.

FT DISULFID 219 228 BY SIMILARITY.

FT DISULFID 224 237 BY SIMILARITY.

FT DISULFID 239 253 BY SIMILARITY.

FT CARBOHYD 271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY).

CC -----

CC SQ SEQUENCE 356 AA; 37795 MW; 29841F097ABE1093 CRC64;

CC -----

CC Query Match 38.6%; Score 1092.5; DB 1; Length 356;

CC Best Local Similarity 63.4%; Pred. No. 1.3e-65;

CC Matches 189; Conservative 30; Mismatches 64; Indels 15; Gaps 6;

CC -----

CC 211 GAVQGHAREAPGANDCSVENGGCHHCAVNPAGPRCCPAGAAQLQADGRSCTASATQSC 270

CC Db 5 GETEGRWREAPGAWACGVVERGCCQHECKGSAGASNCILCPADAAQLQADGRSCGLPAEHPC 64

CC 271 NDLCEHFCVFNPDQGSYSCHCETGYRLAAQDHRCEVDVDDCILPEPSCPQCVNTQGGFE 330

Db 65 HOLCEHFC--HLHGLGNYTCICEAGYQLAADQHRCEVDVDCALQPLSPQRCVNTGGFQ 122  
Qy 331 CHCYPNTDVGECVPEVDPCFRANCEYQCOPLQNTSVCYCAEGFAPIPHEPHRCQMFC 390  
Db 123 CHCTGTVELDVGECVDPVDFDNNCEYQCCQVGRSEHKICACGAFAPVGPAPHKQMFC 182  
Qy 391 NOTACPADCPNTQASCECPGEGYILDGFICTDIDCEGNGFCGVCNHLPGTPECICGP 450  
Db 183 NOTSCPADCPDHPYTCRCPEGYIDEGSTCTDINECDT-NICFGQCHNLPSTVEICICGP 241  
Qy 451 DSALVRHIGTDCSGKV-----DGDSSGSGPPSPPTGSLTLP-PA-VGLVHSG 498  
Db 242 DSALSGQIGIDCDPTQVNEERGTPEDYG--GSGEPVSPPTGATARPSPAPGLHSG 297

RESULT 4  
ID\_CD93 HUMAN STANDARD; PRT; 652 AA.  
AC QNPY3; Q00274;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Complement component, C1q receptor precursor (Complement component 1, q  
DE subcomponent, receptor 1) (C1qR) (C1qR(p)) (C1q/MBL/SPA receptor)  
DE (CD93 antigen) (CDW93).  
GN C1QRI OR CD93.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=9719258; PubMed=9047234;  
RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;  
RT "cDNA cloning and primary structure analysis of C1qR(P), the human  
C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";  
RL Immunity 6:119-129(1997).  
RN [2]  
SEQUENCE FROM N.A., AND VARIANT ALA-318.  
RX MEDLINE=21640567; PubMed=11781389;  
RA Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,  
RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;  
RT "Identification of human CD93 as the phagocytic C1q receptor (C1qR)  
by expression cloning.";  
RL J. Leukoc. Biol. 71:133-140(2002).  
RN [3]  
SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehaeslaio M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mulligan J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.,

RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [4]  
SEQUENCE FROM N.A.  
RX TISSUE=Leukocyte;  
MEDLINE=22389257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
CHARACTERIZATION.  
RX MEDLINE=21990337; PubMed=11994479;  
RA McGreal E.P., Ikewaki N., Akatsu H., Morgan B.P., Gasque P.;  
RT "Human C1qR is identical with CD93 and the mmi-11 antigen but does  
not bind C1q.";  
RL J. Immunol. 168:5222-5232(2002).  
RN [6]  
O-GLYCOSYLATION.  
RX MEDLINE=99192777; PubMed=10092817;  
RA Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;  
RT "C1qR is a heavily O-glycosylated cell surface protein involved in  
the regulation of phagocytic activity.";  
RL J. Immunol. 162:3583-3589(1999).  
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for  
C1q, mannose-binding lectin (MBL2) and pulmonary surfactant  
protein A (SPA). May mediate the enhancement of phagocytosis in  
monocytes and macrophages upon interaction with soluble defense  
collagens. May play a role in intercellular adhesion.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,  
platelets, cells of myeloid origin, such as monocytes and  
neutrophils. Not expressed in cells of lymphoid origin.  
CC -!- PTM: N- and O-glycosylated.  
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
CC -!- SIMILARITY: Contains 5 EGF-like domains.  
CC -!- CAUTION: Has been sometimes referred to as a collectin receptor.  
CC -!- CAUTION: According to Ref.5, C1q is not a ligand for C1QRI.  
CC -!- DATABASE: NAME=PRO; NOTE=PROW 3.1-6(2001);  
WWW="http://www.ncbi.nlm.nih.gov/prov/guide/467246456.g.htm".  
CC -----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL; U94333; AAB53110.1; --  
DR EMBL; AL118508; CAC00597.1; --  
DR EMBL; BC028075; AAB28075.1; --  
DR HSSP; P35555; IEMN.  
DR MIM; 120577; --  
DR GO; GO:0016021; C:integral to membrane; IC.  
DR GO; GO:0004872; F:receptor activity; NAS.

DR GO: 0016337; P: cell-cell adhesion; IDA.  
 DR GO: 004216; P: macrophage activation; NAS.  
 DR GO: 0006909; P: phagocytosis; NAS.  
 DR InterPro: IPR00152; ASx hydrolase\_s.  
 DR InterPro: IPR01881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR Pfam: PF00008; EGF; 5.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00179; EGF\_Ca; 3.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 3.  
 DR PROSITE: PS00615; C-TYPE LECTIN\_1; FALSE\_NEG.  
 DR PROSITE: PS00041; C-TYPE LECTIN\_2; 1.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 DR PROSITE: PS00026; EGF\_3; 3.  
 DR PROSITE: PS01187; EGF\_Ca; 3.  
 KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;  
 KW EGF-like domain; Lectin; Glycoprotein; Polymorphism.  
 FT SIGNAL 1 21  
 FT CHAIN 22 652  
 FT DOMAIN 24 580  
 FT TRANSMEM 581 601  
 FT DOMAIN 602 652  
 FT DOMAIN 32 174  
 FT DOMAIN 260 301  
 FT DOMAIN 302 344  
 FT DOMAIN 345 384  
 FT DOMAIN 385 426  
 FT DOMAIN 427 468  
 FT DOMAIN 594 601  
 FT DISULFID 264 275  
 FT DISULFID 271 285  
 FT DISULFID 287 300  
 FT DISULFID 306 317  
 FT DISULFID 311 328  
 FT DISULFID 330 343  
 FT DISULFID 349 358  
 FT DISULFID 354 367  
 FT DISULFID 369 383  
 FT DISULFID 389 400  
 FT DISULFID 396 409  
 FT DISULFID 411 425  
 FT DISULFID 431 443  
 FT DISULFID 439 452  
 FT DISULFID 454 467  
 FT CARBOHYD 325 325  
 FT VARIANT 318 318  
 FT  
 FT CONFLICT 22 22  
 FT CONFLICT 36 36  
 FT CONFLICT 38 39  
 FT CONFLICT 155 155  
 FT CONFLICT 186 186  
 FT CONFLICT 492 492  
 FT CONFLICT 496 496  
 FT CONFLICT 504 504  
 FT CONFLICT 541 541  
 FT SEQUENCE 652 AA; 68560 MW; EECA0FEAC55FCAC2 CRC64;  
 Query Match 20.4%; Score 578; DB 1; Length 652;  
 Best Local Similarity 31.6%; Pred. No. 2.7e-31;  
 Matches 175; Conservative 52; Mismatches 201; Indels 126; Gaps 31;  
 QY 6 QFGSGQ-----CVEHDCFALYFGPATFLNASQICDGLRGLMTVRSSVAAD---VI 53  
 DB 17 QPGAGTGADTEAVCVGTACTYAHSGKLSAARQAQHCNQNGNLATVSKKEAQHVQRVL 76  
 QY 54 SILLINGDGVGR--RLWIGLQLPFG-CGDKRLGFLRGFWVTGDNNTSYSRWARLDLN 110  
 DB 77 AQLRLREALTARMSKFWIGLQREKGLCDPSL--PLKGFVWVGGEITPYSNWHKELRN 134  
 QY 111 GAPLCGPLCVA--VSAAEATVPE-PIWEQQC-----EVKADGFLCEHFHPATCRPLA 161

Db 135 SC--ISKRCVSLLDLSDQLPLPSRLPKWSEGPCSPGSGNIEGFVCKFSFKGMCPLA 192  
 QY 162 V-ERGAATAAVSITYGTPFAARGADFOALPVGSSAAVAPLGLQLMC-TAPPGAQGHV-- 217  
 Db 193 LGGPG-----QVYITPTFTSSLEAVPASAANVA-----CGEGDKDETQSHYFL 239  
 QY 218 -AREAPGAWD-----CSVENGGCGHAC-NAIPGAPRCOCAPGAALQADGRS 261  
 Db 240 CKEKAPDVEDWSSGSLCVSPKYGCFNNGCGCHQDCEFGDGGFLCGRFGRLDLDT 299  
 QY 262 CTASATQSCNDLCE--HFCVNPDPQGSYSQCMCTGYRLAADQHRCDVDDCILEPSPCP 319  
 Db 300 C-ASRNPCSSSPCRGGATCVLGF-HGKNYTCRCFQGVQLDSSQLDCCVDVDEC--QDSPCA 355  
 QY 320 QRCVNTQGGFECHCYNYDLVDGECVEPDPFCFRANCEYCCQPLNQTSLYLCVCAEGFAP 379  
 Db 356 QECVNTPGGFRCEWGVY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395  
 QY 380 PHEPHRCQMPFCNOTACPADCDPNTQAS--CECPEGYIL--DDGFICTDIDEC--ENGFC 433  
 Db 396 -----CAQGC-TNTDGSFHCSEEGYVLAGEGTQCDVDECVGPGPLC 439  
 QY 434 SGVCHNLPGTFECICGPDSSALVRHIGTDCDSKV-----DGGDSGSGE-----P 477  
 Db 440 DSLCFNTQGSFHCGLPGWVLPFN-GVSCMTGPFVSLGPPSPGPDDEDKEGEGSTVPRAA 498  
 QY 478 PPSPTPGSTLTTPA 491  
 Db 499 TASPTGPGSTPKA 512  
 RESULT 5  
 CD93\_MOUSE STANDARD; PRT; 644 AA.  
 AC CD93\_MOUSE  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Complement component C1q receptor precursor (Complement component 1, q  
 subcomponent, receptor 1) (ClqRp) (ClqRp) (ClqRp/SPA receptor)  
 DE (CD93 antigen) (Cell surface antigen AA4) (Lymphocyte antigen 68).  
 GN CLQRI OR CD93 OR CLQRP OR LY68 OR AA4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=20528605; PubMed=11074255;  
 RA Kim T.S., Park M., Nepomuceno R.R., Palmirani G., Winokur S.,  
 RA Cotman C.A., Bengtsson U., Tenner A.J.;  
 RT "Characterization of the murine homolog of ClqRp: identical cellular  
 expression pattern, chromosomal location and functional activity of  
 the human and murine ClqRp(P).";  
 RL Mol. Immunol. 37:377-389(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukemia;  
 RX MEDLINE=99330438; PubMed=10403644;  
 RA Petrenko O., Beavis A., Klaine M., Kittappa R., Godin I.,  
 RA Lemischka I.R.;  
 RT "The molecular characterization of the fetal stem cell marker AA4";  
 RL Immunity 10:691-700(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv; TISSUE=Endothelial cells, and Spleen;  
 RX MEDLINE=99359842; PubMed=10430665;  
 RA Norworthy P.J., Taylor P.R., Walport M.J., Botta M.;  
 RT "Cloning of the mouse homolog of the 126-kDa human Clq/MBL/SP-A  
 receptor, ClqRp.";  
 RL Mamm. Genome 10:789-793(1999).







OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PVG; TISSUE=Natural killer cells;  
 RX MEDLINE=20545218; PubMed=11093152;  
 RA Lovik G., Vaage J.F., Dissen E., Spierer C., Ryan J.C., Rolstad B.;  
 RT "Characterization and molecular cloning of rat C1qR, a receptor on NK  
 cells.";  
 RL Eur. J. Immunol. 30:3355-3362(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Lung;  
 RX MEDLINE=20507883; PubMed=10334210;  
 RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;  
 RT "Molecular and cellular properties of the rat A4 antigen, a C-type  
 lectin-like receptor with structural homology to thrombomodulin.";  
 RL J. Biol. Chem. 275:34382-34392(2000).  
 CC -|- FUNCTION: Receptor (or element of a larger receptor complex) for  
 C1q, mannose-binding lectin (MBL2) and pulmonary surfactant  
 protein A (SPA). May mediate the enhancement of phagocytosis in  
 monocytes and macrophages upon interaction with soluble defense  
 collagens. May play a role in intercellular adhesion.  
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -|- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and  
 heart. Expressed at lower level in brain, thymus, liver, spleen,  
 intestine, kidney, adrenal gland, muscle and testis. Expressed on  
 endothelial cells, platelets, undifferentiated monocytes and  
 circulating natural killer cells.  
 CC -|- PTM: N- and O-glycosylated (By similarity).  
 CC -|- SIMILARITY: Contains 1 C-type lectin family domain.  
 CC -|- SIMILARITY: Contains 5 EGF-like domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF136537; AAG01572.1; -;  
 CC EMBL; AF160978; AAF80402.1; -;  
 CC HSP; P35555; IEMN  
 CC GO; GO:0016021; C: integral to membrane; ISS.  
 CC GO; GO:0004872; F: receptor activity; ISS.  
 CC GO; GO:0016337; P: cell-cell adhesion; ISS.  
 CC GO; GO:0042116; P: macrophage activation; ISS.  
 CC GO; GO:0006909; P: phagocytosis; ISS.  
 CC InterPro; IPR000152; Asx hydroxyl S.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR006209; EGF-like.  
 CC InterPro; IPR001304; Lectin\_C.  
 CC Pfam; PF00008; EGF; 4.  
 CC SMART; SM00034; LECTC1; 1.  
 CC SMART; SM00179; EGF\_CA; 3.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 3.  
 CC PROSITE; PS00615; C\_TYPE\_LECTIN\_1; FALSE\_NEG.  
 CC PROSITE; PS00041; C\_TYPE\_LECTIN\_2; 1.  
 CC PROSITE; PS01186; EGF\_2; 3.  
 CC PROSITE; PS00026; EGF\_3; 4.  
 CC PROSITE; PS01187; EGF\_CA; 3.  
 CC Cell adhesion; Receptor; Repeat; Signal; Transmembrane;  
 KW EGF-like domain; Lectin; Glycoprotein.  
 KW SIGNAL  
 FT SIGNAL 1 23  
 FT CHAIN 24 643  
 FT DOMAIN 24 571  
 FT TRANSMEM 572 592  
 FT DOMAIN 593 643  
 FT DOMAIN 31 173  
 FT DOMAIN 257 298  
 FT DOMAIN 299 341  
 FT DOMAIN 342 381  
 FT DOMAIN 381 445  
 FT DOMAIN 445 485  
 FT DOMAIN 485 516  
 FT DOMAIN 516 547  
 FT DOMAIN 547 578  
 FT DOMAIN 578 609  
 FT DOMAIN 609 640  
 FT DOMAIN 640 671  
 FT DOMAIN 671 702  
 FT DOMAIN 702 733  
 FT DOMAIN 733 764  
 FT DOMAIN 764 795  
 FT DOMAIN 795 826  
 FT DOMAIN 826 857  
 FT DOMAIN 857 888  
 FT DOMAIN 888 919  
 FT DOMAIN 919 950  
 FT DOMAIN 950 981  
 FT DOMAIN 981 1012  
 FT DOMAIN 1012 1043  
 FT DOMAIN 1043 1074  
 FT DOMAIN 1074 1105  
 FT DOMAIN 1105 1136  
 FT DOMAIN 1136 1167  
 FT DOMAIN 1167 1198  
 FT DOMAIN 1198 1229  
 FT DOMAIN 1229 1260  
 FT DOMAIN 1260 1291  
 FT DOMAIN 1291 1322  
 FT DOMAIN 1322 1353  
 FT DOMAIN 1353 1384  
 FT DOMAIN 1384 1415  
 FT DOMAIN 1415 1446  
 FT DOMAIN 1446 1477  
 FT DOMAIN 1477 1508  
 FT DOMAIN 1508 1539  
 FT DOMAIN 1539 1570  
 FT DOMAIN 1570 1601  
 FT DOMAIN 1601 1632  
 FT DOMAIN 1632 1663  
 FT DOMAIN 1663 1694  
 FT DOMAIN 1694 1725  
 FT DOMAIN 1725 1756  
 FT DOMAIN 1756 1787  
 FT DOMAIN 1787 1818  
 FT DOMAIN 1818 1849  
 FT DOMAIN 1849 1880  
 FT DOMAIN 1880 1911  
 FT DOMAIN 1911 1942  
 FT DOMAIN 1942 1973  
 FT DOMAIN 1973 2004  
 FT DOMAIN 2004 2035  
 FT DOMAIN 2035 2066  
 FT DOMAIN 2066 2097  
 FT DOMAIN 2097 2128  
 FT DOMAIN 2128 2159  
 FT DOMAIN 2159 2190  
 FT DOMAIN 2190 2221  
 FT DOMAIN 2221 2252  
 FT DOMAIN 2252 2283  
 FT DOMAIN 2283 2314  
 FT DOMAIN 2314 2345  
 FT DOMAIN 2345 2376  
 FT DOMAIN 2376 2407  
 FT DOMAIN 2407 2438  
 FT DOMAIN 2438 2469  
 FT DOMAIN 2469 2500  
 FT DOMAIN 2500 2531  
 FT DOMAIN 2531 2562  
 FT DOMAIN 2562 2593  
 FT DOMAIN 2593 2624  
 FT DOMAIN 2624 2655  
 FT DOMAIN 2655 2686  
 FT DOMAIN 2686 2717  
 FT DOMAIN 2717 2748  
 FT DOMAIN 2748 2779  
 FT DOMAIN 2779 2810  
 FT DOMAIN 2810 2841  
 FT DOMAIN 2841 2872  
 FT DOMAIN 2872 2903  
 FT DOMAIN 2903 2934  
 FT DOMAIN 2934 2965  
 FT DOMAIN 2965 2996  
 FT DOMAIN 2996 3027  
 FT DOMAIN 3027 3058  
 FT DOMAIN 3058 3089  
 FT DOMAIN 3089 3120  
 FT DOMAIN 3120 3151  
 FT DOMAIN 3151 3182  
 FT DOMAIN 3182 3213  
 FT DOMAIN 3213 3244  
 FT DOMAIN 3244 3275  
 FT DOMAIN 3275 3306  
 FT DOMAIN 3306 3337  
 FT DOMAIN 3337 3368  
 FT DOMAIN 3368 3399  
 FT DOMAIN 3399 3430  
 FT DOMAIN 3430 3461  
 FT DOMAIN 3461 3492  
 FT DOMAIN 3492 3523  
 FT DOMAIN 3523 3554  
 FT DOMAIN 3554 3585  
 FT DOMAIN 3585 3616  
 FT DOMAIN 3616 3647  
 FT DOMAIN 3647 3678  
 FT DOMAIN 3678 3709  
 FT DOMAIN 3709 3740  
 FT DOMAIN 3740 3771  
 FT DOMAIN 3771 3802  
 FT DOMAIN 3802 3833  
 FT DOMAIN 3833 3864  
 FT DOMAIN 3864 3895  
 FT DOMAIN 3895 3926  
 FT DOMAIN 3926 3957  
 FT DOMAIN 3957 3988  
 FT DOMAIN 3988 4019  
 FT DOMAIN 4019 4050  
 FT DOMAIN 4050 4081  
 FT DOMAIN 4081 4112  
 FT DOMAIN 4112 4143  
 FT DOMAIN 4143 4174  
 FT DOMAIN 4174 4205  
 FT DOMAIN 4205 4236  
 FT DOMAIN 4236 4267  
 FT DOMAIN 4267 4298  
 FT DOMAIN 4298 4329  
 FT DOMAIN 4329 4360  
 FT DOMAIN 4360 4391  
 FT DOMAIN 4391 4422  
 FT DOMAIN 4422 4453  
 FT DOMAIN 4453 4484  
 FT DOMAIN 4484 4515  
 FT DOMAIN 4515 4546  
 FT DOMAIN 4546 4577  
 FT DOMAIN 4577 4608  
 FT DOMAIN 4608 4639  
 FT DOMAIN 4639 4670  
 FT DOMAIN 4670 4701  
 FT DOMAIN 4701 4732  
 FT DOMAIN 4732 4763  
 FT DOMAIN 4763 4794  
 FT DOMAIN 4794 4825  
 FT DOMAIN 4825 4856  
 FT DOMAIN 4856 4887  
 FT DOMAIN 4887 4918  
 FT DOMAIN 4918 4949  
 FT DOMAIN 4949 4980  
 FT DOMAIN 4980 5011  
 FT DOMAIN 5011 5042  
 FT DOMAIN 5042 5073  
 FT DOMAIN 5073 5104  
 FT DOMAIN 5104 5135  
 FT DOMAIN 5135 5166  
 FT DOMAIN 5166 5197  
 FT DOMAIN 5197 5228  
 FT DOMAIN 5228 5259  
 FT DOMAIN 5259 5290  
 FT DOMAIN 5290 5321  
 FT DOMAIN 5321 5352  
 FT DOMAIN 5352 5383  
 FT DOMAIN 5383 5414  
 FT DOMAIN 5414 5445  
 FT DOMAIN 5445 5476  
 FT DOMAIN 5476 5507  
 FT DOMAIN 5507 5538  
 FT DOMAIN 5538 5569  
 FT DOMAIN 5569 5600  
 FT DOMAIN 5600 5631  
 FT DOMAIN 5631 5662  
 FT DOMAIN 5662 5693  
 FT DOMAIN 5693 5724  
 FT DOMAIN 5724 5755  
 FT DOMAIN 5755 5786  
 FT DOMAIN 5786 5817  
 FT DOMAIN 5817 5848  
 FT DOMAIN 5848 5879  
 FT DOMAIN 5879 5910  
 FT DOMAIN 5910 5941  
 FT DOMAIN 5941 5972  
 FT DOMAIN 5972 6003  
 FT DOMAIN 6003 6034  
 FT DOMAIN 6034 6065  
 FT DOMAIN 6065 6096  
 FT DOMAIN 6096 6127  
 FT DOMAIN 6127 6158  
 FT DOMAIN 6158 6189  
 FT DOMAIN 6189 6220  
 FT DOMAIN 6220 6251  
 FT DOMAIN 6251 6282  
 FT DOMAIN 6282 6313  
 FT DOMAIN 6313 6344  
 FT DOMAIN 6344 6375  
 FT DOMAIN 6375 6406  
 FT DOMAIN 6406 6437  
 FT DOMAIN 6437 6468  
 FT DOMAIN 6468 6499  
 FT DOMAIN 6499 6530  
 FT DOMAIN 6530 6561  
 FT DOMAIN 6561 6592  
 FT DOMAIN 6592 6623  
 FT DOMAIN 6623 6654  
 FT DOMAIN 6654 6685  
 FT DOMAIN 6685 6716  
 FT DOMAIN 6716 6747  
 FT DOMAIN 6747 6778  
 FT DOMAIN 6778 6809  
 FT DOMAIN 6809 6840  
 FT DOMAIN 6840 6871  
 FT DOMAIN 6871 6902  
 FT DOMAIN 6902 6933  
 FT DOMAIN 6933 6964  
 FT DOMAIN 6964 6995  
 FT DOMAIN 6995 7026  
 FT DOMAIN 7026 7057  
 FT DOMAIN 7057 7088  
 FT DOMAIN 7088 7119  
 FT DOMAIN 7119 7150  
 FT DOMAIN 7150 7181  
 FT DOMAIN 7181 7212  
 FT DOMAIN 7212 7243  
 FT DOMAIN 7243 7274  
 FT DOMAIN 7274 7305  
 FT DOMAIN 7305 7336  
 FT DOMAIN 7336 7367  
 FT DOMAIN 7367 7398  
 FT DOMAIN 7398 7429  
 FT DOMAIN 7429 7460  
 FT DOMAIN 7460 7491  
 FT DOMAIN 7491 7522  
 FT DOMAIN 7522 7553  
 FT DOMAIN 7553 7584  
 FT DOMAIN 7584 7615  
 FT DOMAIN 7615 7646  
 FT DOMAIN 7646 7677  
 FT DOMAIN 7677 7708  
 FT DOMAIN 7708 7739  
 FT DOMAIN 7739 7770  
 FT DOMAIN 7770 7801  
 FT DOMAIN 7801 7832  
 FT DOMAIN 7832 7863  
 FT DOMAIN 7863 7894  
 FT DOMAIN 7894 7925  
 FT DOMAIN 7925 7956  
 FT DOMAIN 7956 7987  
 FT DOMAIN 7987 8018  
 FT DOMAIN 8018 8049  
 FT DOMAIN 8049 8080  
 FT DOMAIN 8080 8111  
 FT DOMAIN 8111 8142  
 FT DOMAIN 8142 8173  
 FT DOMAIN 8173 8204  
 FT DOMAIN 8204 8235  
 FT DOMAIN 8235 8266  
 FT DOMAIN 8266 8297  
 FT DOMAIN 8297 8328  
 FT DOMAIN 8328 8359  
 FT DOMAIN 8359 8390  
 FT DOMAIN 8390 8421  
 FT DOMAIN 8421 8452  
 FT DOMAIN 8452 8483  
 FT DOMAIN 8483 8514  
 FT DOMAIN 8514 8545  
 FT DOMAIN 8545 8576  
 FT DOMAIN 8576 8607  
 FT DOMAIN 8607 8638  
 FT DOMAIN 8638 8669  
 FT DOMAIN 8669 8700  
 FT DOMAIN 8700 8731  
 FT DOMAIN 8731 8762  
 FT DOMAIN 8762 8793  
 FT DOMAIN 8793 8824  
 FT DOMAIN 8824 8855  
 FT DOMAIN 8855 8886  
 FT DOMAIN 8886 8917  
 FT DOMAIN 8917 8948  
 FT DOMAIN 8948 8979  
 FT DOMAIN 8979 9010  
 FT DOMAIN 9010 9041  
 FT DOMAIN 9041 9072  
 FT DOMAIN 9072 9103  
 FT DOMAIN 9103 9134  
 FT DOMAIN 9134 9165  
 FT DOMAIN 9165 9196  
 FT DOMAIN 9196 9227  
 FT DOMAIN 9227 9258  
 FT DOMAIN 9258 9289  
 FT DOMAIN 9289 9320  
 FT DOMAIN 9320 9351  
 FT DOMAIN 9351 9382  
 FT DOMAIN 9382 9413  
 FT DOMAIN 9413 9444  
 FT DOMAIN 9444 9475  
 FT DOMAIN 9475 9506  
 FT DOMAIN 9506 9537  
 FT DOMAIN 9537 9568  
 FT DOMAIN 9568 9599  
 FT DOMAIN 9599 9630  
 FT DOMAIN 9630 9661  
 FT DOMAIN 9661 9692  
 FT DOMAIN 9692 9723  
 FT DOMAIN 9723 9754  
 FT DOMAIN 9754 9785  
 FT DOMAIN 9785 9816  
 FT DOMAIN 9816 9847  
 FT DOMAIN 9847 9878  
 FT DOMAIN 9878 9909  
 FT DOMAIN 9909 9940  
 FT DOMAIN 9940 9971  
 FT DOMAIN 9971 10002  
 FT DOMAIN 10002 10033  
 FT DOMAIN 10033 10064  
 FT DOMAIN 10064 10095  
 FT DOMAIN 10095 10126  
 FT DOMAIN 10126 10157  
 FT DOMAIN 10157 10188  
 FT DOMAIN 10188 10219  
 FT DOMAIN 10219 10250  
 FT DOMAIN 10250 10281  
 FT DOMAIN 10281 10312  
 FT DOMAIN 10312 10343  
 FT DOMAIN 10343 10374  
 FT DOMAIN 10374 10405  
 FT DOMAIN 10405 10436  
 FT DOMAIN 10436 10467  
 FT DOMAIN 10467 10498  
 FT DOMAIN 10498 10529  
 FT DOMAIN 10529 10560  
 FT DOMAIN 10560 10591  
 FT DOMAIN 10591 10622  
 FT DOMAIN 10622 10653  
 FT DOMAIN 10653 10684  
 FT DOMAIN 10684 10715  
 FT DOMAIN 10715 10746  
 FT DOMAIN 10746 10777  
 FT DOMAIN 10777 10808  
 FT DOMAIN 10808 10839  
 FT DOMAIN 10839 10870  
 FT DOMAIN 10870 10901  
 FT DOMAIN 10901 10932  
 FT DOMAIN 10932 10963  
 FT DOMAIN 10963 10994  
 FT DOMAIN 10994 11025  
 FT DOMAIN 11025 11056  
 FT DOMAIN 11056 11087  
 FT DOMAIN 11087 11118  
 FT DOMAIN 11118 11149  
 FT DOMAIN 11149 11180  
 FT DOMAIN 11180 11211  
 FT DOMAIN 11211 11242  
 FT DOMAIN 11242 11273  
 FT DOMAIN 11273 11304  
 FT DOMAIN 11304 11335  
 FT DOMAIN 11335 11366  
 FT DOMAIN 11366 11397  
 FT DOMAIN 11397 11428  
 FT DOMAIN 11428 11459  
 FT DOMAIN 11459 11490  
 FT DOMAIN 11490 11521  
 FT DOMAIN 11521 11552  
 FT DOMAIN 11552 11583  
 FT DOMAIN 11583 11614  
 FT DOMAIN 11614 11645  
 FT DOMAIN 11645 11676  
 FT DOMAIN 11676 11707  
 FT DOMAIN 11707 11738  
 FT DOMAIN 11738 11769  
 FT DOMAIN 11769 11800  
 FT DOMAIN 11800 11831  
 FT DOMAIN 11831 11862  
 FT DOMAIN 11862 11893  
 FT DOMAIN 11893 11924  
 FT DOMAIN 11924 11955  
 FT DOMAIN 11955 11986  
 FT DOMAIN 11986 12017  
 FT DOMAIN 12017 12048  
 FT DOMAIN 12048 12079  
 FT DOMAIN 12079 12110  
 FT DOMAIN 12110 12141  
 FT DOMAIN 12141 12172  
 FT DOMAIN 12172 12203  
 FT DOMAIN 12203 12234  
 FT DOMAIN 12234 12265  
 FT DOMAIN 12265 12296  
 FT DOMAIN 12296 12327  
 FT DOMAIN 12327 12358  
 FT DOMAIN 12358 12389  
 FT DOMAIN 12389 12420  
 FT DOMAIN 12420 12451  
 FT DOMAIN 12451 12482  
 FT DOMAIN 12482 12513  
 FT DOMAIN 12513 12544  
 FT DOMAIN 12544 12575  
 FT DOMAIN 12575 12606  
 FT DOMAIN 12606 12637  
 FT DOMAIN 12637 12668  
 FT DOMAIN 12668 12699  
 FT DOMAIN 12699 12730  
 FT DOMAIN 12730 12761  
 FT DOMAIN 12761 12792  
 FT DOMAIN 12792 12823  
 FT DOMAIN 12823 12854  
 FT DOMAIN 12854 12885  
 FT DOMAIN 12885 12916  
 FT DOMAIN 12916 12947  
 FT DOMAIN 12947 12978  
 FT DOMAIN 12978 13009  
 FT DOMAIN 13009 13040  
 FT DOMAIN 13040 13071  
 FT DOMAIN 13071 13102  
 FT DOMAIN 13102 13133  
 FT DOMAIN 13133 13164  
 FT DOMAIN 13164 13195  
 FT DOMAIN 13195 13226  
 FT DOMAIN 13226 13257  
 FT DOMAIN 13257 13288  
 FT DOMAIN 13288 13319  
 FT DOMAIN 13319 13350  
 FT DOMAIN 13350 13381  
 FT DOMAIN 13381 13412  
 FT DOMAIN 13412 13443  
 FT DOMAIN 13443 13474  
 FT DOMAIN 13474 13505  
 FT DOMAIN 13505 13536  
 FT DOMAIN 13536 13567  
 FT DOMAIN 13567 13598  
 FT DOMAIN 13598 13629  
 FT DOMAIN 13629 13660  
 FT DOMAIN 13660 13691  
 FT DOMAIN 13691 13722  
 FT DOMAIN 13722 13753  
 FT DOMAIN 13753 13784  
 FT DOMAIN 13784 13815  
 FT DOMAIN 13815 13846  
 FT DOMAIN 13846 13877  
 FT DOMAIN 13877 13908  
 FT DOMAIN 13908 13939  
 FT DOMAIN 13939 13970  
 FT DOMAIN 13970 14001  
 FT DOMAIN 14001 14032  
 FT DOMAIN 14032 14063  
 FT DOMAIN 14063 14094  
 FT DOMAIN 14094 14125  
 FT DOMAIN 14125 14156  
 FT DOMAIN 14156 14187  
 FT DOMAIN 14187 14218  
 FT DOMAIN 14218 14249  
 FT DOMAIN 14249 14280  
 FT DOMAIN 14280 14311  
 FT DOMAIN 14311 14342  
 FT DOMAIN 14342 14373  
 FT DOMAIN 14373 14404  
 FT DOMAIN 14404 14435  
 FT DOMAIN 14435 14466  
 FT DOMAIN 14466 14497  
 FT DOMAIN 14497 14528  
 FT DOMAIN 14528 14559  
 FT DOMAIN 14559 14590  
 FT DOMAIN 14590 14621  
 FT DOMAIN 14621 14652  
 FT DOMAIN 14652 14683  
 FT DOMAIN 14683 14714  
 FT DOMAIN 14714 14745  
 FT DOMAIN 14745 14776  
 FT DOMAIN 14776 14807  
 FT DOMAIN 14807 14838  
 FT DOMAIN 14838 14869  
 FT DOMAIN 14869 14900  
 FT DOMAIN 14900 14931  
 FT DOMAIN 14931 14962  
 FT DOMAIN 14962 14993  
 FT DOMAIN 14993 15024  
 FT DOMAIN 15024 15055  
 FT DOMAIN 15055 15086  
 FT DOMAIN 15086 15117  
 FT DOMAIN 15117 15148  
 FT DOMAIN 15148 15179  
 FT DOMAIN 15179 15210  
 FT DOMAIN 15210 15241  
 FT DOMAIN 15241 15272  
 FT DOMAIN 15272 15303  
 FT DOMAIN 15303 15334  
 FT DOMAIN 15334 15365  
 FT DOMAIN 15365 15396  
 FT DOMAIN 15396 15427  
 FT DOMAIN 15427 15458  
 FT DOMAIN 15458 15489  
 FT DOMAIN 15489 15520  
 FT DOMAIN 15520 15551  
 FT DOMAIN 15551 15582  
 FT DOMAIN 15582 15613  
 FT DOMAIN 15613 15644  
 FT DOMAIN 15644 15675  
 FT DOMAIN 15675 15706  
 FT DOMAIN 15706 15737  
 FT DOMAIN 15737 15768  
 FT DOMAIN 15768 15799  
 FT DOMAIN 15799 15830  
 FT DOMAIN 15830 15861  
 FT DOMAIN 15861 15892  
 FT DOMAIN 15892 15923  
 FT DOMAIN 15923 15954  
 FT DOMAIN 15954 15985  
 FT DOMAIN 15985 16016  
 FT DOMAIN 16016 16047  
 FT DOMAIN 16047 16078  
 FT DOMAIN 16078 16109  
 FT DOMAIN 16109 16140  
 FT DOMAIN 16140 16171  
 FT DOMAIN 16171 16202  
 FT DOMAIN 16202 16233  
 FT DOMAIN 16233 16264  
 FT DOMAIN 16264 16295  
 FT DOMAIN 16295 16326  
 FT DOMAIN 16326 16357  
 FT DOMAIN 16357 16388  
 FT DOMAIN 16388 16419  
 FT DOMAIN 16419 16450  
 FT DOMAIN 16450 16481  
 FT DOMAIN 16481 16512  
 FT DOMAIN 16512 16543  
 FT DOMAIN 16543 16574  
 FT DOMAIN 16574 16605  
 FT DOMAIN 16605 16636  
 FT DOMAIN 16636 16667  
 FT DOMAIN 16667 16698  
 FT DOMAIN 16698 16729  
 FT DOMAIN 16729 16760  
 FT DOMAIN 16760 16791  
 FT DOMAIN 16791 16822  
 FT DOMAIN 16822 16853  
 FT DOMAIN 16853 16884  
 FT DOMAIN 16884 16915  
 FT DOMAIN 16915 16946  
 FT DOMAIN 16946 16977  
 FT DOMAIN 16977 17008  
 FT DOMAIN 17008 17039  
 FT DOMAIN 17039 17070  
 FT DOMAIN 17070 17101  
 FT DOMAIN 17101 17132  
 FT DOMAIN 17132 17163  
 FT DOMAIN 17163 17194  
 FT DOMAIN 17194 17225  
 FT DOMAIN 17225 17256  
 FT DOMAIN 17256 17287  
 FT DOMAIN 17287 17318  
 FT DOMAIN 17318 17349  
 FT DOMAIN 17349 17380  
 FT DOMAIN 17380 17411  
 FT DOMAIN 17411 17442  
 FT DOMAIN 17442 17473  
 FT DOMAIN 17473 17504  
 FT DOMAIN 17504 17535  
 FT DOMAIN 17535 17566  
 FT DOMAIN 17566 17597  
 FT DOMAIN 17597 17628  
 FT DOMAIN 17628 17659  
 FT DOMAIN 17659 17690  
 FT DOMAIN 17690 17721  
 FT DOMAIN 17721 17752  
 FT DOMAIN 17752 17783  
 FT DOMAIN 17783 17814  
 FT DOMAIN 17814 17845  
 FT DOMAIN 17845 17876  
 FT DOMAIN 17876 17907  
 FT DOMAIN 17907 17938  
 FT DOMAIN 17938 17969  
 FT DOMAIN 17969 17999  
 FT DOMAIN 17999 18030  
 FT DOMAIN 18030 18061  
 FT DOMAIN 18061 18092  
 FT DOMAIN 18092 18123  
 FT DOMAIN 18123 18154  
 FT DOMAIN 18154 18185  
 FT DOMAIN 18185 18216  
 FT DOMAIN 18216 18247  
 FT DOMAIN 18247 18278  
 FT DOMAIN 18278 18309  
 FT DOMAIN 18309 18340  
 FT DOMAIN 18340 18371  
 FT DOMAIN 18371 18402  
 FT DOMAIN 18402 18433  
 FT DOMAIN 18433 18464  
 FT DOMAIN 18464 18495  
 FT DOMAIN 18495 18526  
 FT DOMAIN 18526 18557  
 FT DOMAIN 18557 18588  
 FT DOMAIN 18588 18619  
 FT DOMAIN 18619 18650  
 FT DOMAIN 18650 18681  
 FT DOMAIN 18681 18712  
 FT DOMAIN 18712 18743  
 FT DOMAIN 18743 18774  
 FT DOMAIN 18774 18805  
 FT DOMAIN 18805 18836  
 FT DOMAIN 18836 18867  
 FT DOMAIN 18867 18898  
 FT DOMAIN 18898 18929  
 FT DOMAIN 18929 18960  
 FT DOMAIN 18960 18991  
 FT DOMAIN 18991 19022  
 FT DOMAIN 19022 19053  
 FT DOMAIN 19053 19084  
 FT DOMAIN 19084 19115  
 FT DOMAIN 19115 19146  
 FT DOMAIN 19146 19177  
 FT DOMAIN 19177 19208  
 FT DOMAIN 19208 19239  
 FT DOMAIN 19239 19270  
 FT DOMAIN 19270 19301  
 FT DOMAIN 19301 19332  
 FT DOMAIN 19332 19363  
 FT DOMAIN 19363 19394  
 FT DOMAIN 19394 19425  
 FT DOMAIN 19425 19456  
 FT DOMAIN 19456 19487  
 FT DOMAIN 19487 19518  
 FT DOMAIN 19518 19549  
 FT DOMAIN 19549 19580  
 FT DOMAIN 19580 19611  
 FT DOMAIN 19611 19642  
 FT DOMAIN 19642 19673  
 FT DOMAIN 19673 19704  
 FT DOMAIN 19704 19735  
 FT DOMAIN 19735 19766  
 FT DOMAIN 19766 19797  
 FT DOMAIN 19797 19828  
 FT DOMAIN 19828 19859  
 FT DOMAIN 19859 19890  
 FT DOMAIN 19890 19921  
 FT DOMAIN 19921 19952  
 FT DOMAIN 19952 19983  
 FT DOMAIN 19983 20014  
 FT DOMAIN 20014 20045  
 FT DOMAIN 20045 20076  
 FT DOMAIN 20076

OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RX MEDLINE=95263670; PubMed=7744963;  
 RA Zhang H., Hu W., Ramirez F.;  
 RT "Developmental expression of fibrillin genes suggests heterogeneity  
 of extracellular microfibrils.";  
 RL J. Cell Biol. 129:1165-1176(1995).  
 RN [2]  
 RP SEQUENCE OF 210-317 FROM N.A.  
 RX MEDLINE=94140368; PubMed=8307578;  
 RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,  
 RA Francke U.;  
 RT mouse chromosomes 2 and 18";  
 RL Genomics 18:667-672(1993).  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 that binds calcium. Fibrillin-2-containing microfibrils regulate  
 the early process of elastic fiber assembly.  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; L39790; AAA74908.1; -  
 DR EMBL; S69359; AAC60685.1; -  
 DR PIR; A57278; A57278.  
 DR HGSP; P35555; 1EMN.  
 DR MGD; PGI:95490; Fbn2.  
 DR GO; GO:0030326; P:limb morphogenesis; IMP.  
 DR InterPro; IPR000152; ASX Hydroxyl\_S.  
 DR InterPro; IPR001881; EGF CA.  
 DR InterPro; IPR001438; EGF II.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR002212; Fibril-assoc.  
 DR Pfam; PF00008; EGF; 44.  
 DR Pfam; PF00683; TB; 9.  
 DR PRINTS; PR00010; EGFBL00D.  
 DR SMART; SM00179; EGF CA; 43.  
 DR PROSITE; PS00010; ASX HYDROXYL; 43.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 36.  
 DR PROSITE; PS00026; EGF\_3; 45.  
 DR PROSITE; PS01187; EGF\_CA; 43.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 Repeat; Signal; Multigene family.  
 FT SIGNAL 1 28  
 FT CHAIN 29 2907 POTENTIAL.  
 FT DOMAIN 111 142 FIBRILLIN 2.  
 FT EGF-LIKE 1.  
 FT DOMAIN 145 176 EGF-LIKE 2.  
 FT DOMAIN 176 208 EGF-LIKE 3.  
 FT DOMAIN 276 317 EGF-LIKE 4.  
 FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.  
 FT REPEAT 360 426 EGF-LIKE 6, CALCIUM-BINDING.  
 FT DOMAIN 487 527 TGFBP 1.  
 FT EGF-LIKE 7.  
 FT DOMAIN 528 567 EGF-LIKE 8, CALCIUM-BINDING.  
 FT DOMAIN 568 609 EGF-LIKE 9, CALCIUM-BINDING.  
 FT DOMAIN 610 650 EGF-LIKE 10, CALCIUM-BINDING.  
 FT DOMAIN 651 691 EGF-LIKE 11, CALCIUM-BINDING.  
 FT REPEAT 692 760 TGFBP 2.  
 FT DOMAIN 761 802 EGF-LIKE 12, CALCIUM-BINDING.  
 FT DOMAIN 803 844 EGF-LIKE 13, CALCIUM-BINDING.  
 FT DOMAIN 845 883 EGF-LIKE 14, CALCIUM-BINDING.  
 FT DOMAIN 948 989 TGFBP 3.  
 FT REPEAT 990 1065

1066 1107 1150 1192 1235 1275 1317 1359 1400 1441 1483 1524 1565 1642 1684 1726 1768 1800 1842 1884 1926 1965 2008 2048 2090 2163 2205 2245 2286 2327 2372 2411 2442 2484 2524 2563 2606 2646 2687 2727 2768 2807 2847 2888 2927 3011 3051 3091 3131 3171 3211 3251 3291 3331 3371 3411 3451 3491 3531 3571 3611 3651 3691 3731 3771 3811 3851 3891 3931 3971 4011 4051 4091 4131 4171 4211 4251 4291 4331 4371 4411 4451 4491 4531 4571 4611 4651 4691 4731 4771 4811 4851 4891 4931 4971 5011 5051 5091 5131 5171 5211 5251 5291 5331 5371 5411 5451 5491 5531 5571 5611 5651 5691 5731 5771 5811 5851 5891 5931 5971 6011 6051 6091 6131 6171 6211 6251 6291 6331 6371 6411 6451 6491 6531 6571 6611 6651 6691 6731 6771 6811 6851 6891 6931 6971 7011 7051 7091 7131 7171 7211 7251 7291 7331 7371 7411 7451 7491 7531 7571 7611 7651 7691 7731 7771 7811 7851 7891 7931 7971 8011 8051 8091 8131 8171 8211 8251 8291 8331 8371 8411 8451 8491 8531 8571 8611 8651 8691 8731 8771 8811 8851 8891 8931 8971 9011 9051 9091 9131 9171 9211 9251 9291 9331 9371 9411 9451 9491 9531 9571 9611 9651 9691 9731 9771 9811 9851 9891 9931 9971 10011 10051 10091 10131 10171 10211 10251 10291 10331 10371 10411 10451 10491 10531 10571 10611 10651 10691 10731 10771 10811 10851 10891 10931 10971 11011 11051 11091 11131 11171 11211 11251 11291 11331 11371 11411 11451 11491 11531 11571 11611 11651 11691 11731 11771 11811 11851 11891 11931 11971 12011 12051 12091 12131 12171 12211 12251 12291 12331 12371 12411 12451 12491 12531 12571 12611 12651 12691 12731 12771 12811 12851 12891 12931 12971 13011 13051 13091 13131 13171 13211 13251 13291 13331 13371 13411 13451 13491 13531 13571 13611 13651 13691 13731 13771 13811 13851 13891 13931 13971 14011 14051 14091 14131 14171 14211 14251 14291 14331 14371 14411 14451 14491 14531 14571 14611 14651 14691 14731 14771 14811 14851 14891 14931 14971 15011 15051 15091 15131 15171 15211 15251 15291 15331 15371 15411 15451 15491 15531 15571 15611 15651 15691 15731 15771 15811 15851 15891 15931 15971 16011 16051 16091 16131 16171 16211 16251 16291 16331 16371 16411 16451 16491 16531 16571 16611 16651 16691 16731 16771 16811 16851 16891 16931 16971 17011 17051 17091 17131 17171 17211 17251 17291 17331 17371 17411 17451 17491 17531 17571 17611 17651 17691 17731 17771 17811 17851 17891 17931 17971 18011 18051 18091 18131 18171 18211 18251 18291 18331 18371 18411 18451 18491 18531 18571 18611 18651 18691 18731 18771 18811 18851 18891 18931 18971 19011 19051 19091 19131 19171 19211 19251 19291 19331 19371 19411 19451 19491 19531 19571 19611 19651 19691 19731 19771 19811 19851 19891 19931 19971 20011 20051 20091 20131 20171 20211 20251 20291 20331 20371 20411 20451 20491 20531 20571 20611 20651 20691 20731 20771 20811 20851 20891 20931 20971 21011 21051 21091 21131 21171 21211 21251 21291 21331 21371 21411 21451 21491 21531 21571 21611 21651 21691 21731 21771 21811 21851 21891 21931 21971 22011 22051 22091 22131 22171 22211 22251 22291 22331 22371 22411 22451 22491 22531 22571 22611 22651 22691 22731 22771 22811 22851 22891 22931 22971 23011 23051 23091 23131 23171 23211 23251 23291 23331 23371 23411 23451 23491 23531 23571 23611 23651 23691 23731 23771 23811 23851 23891 23931 23971 24011 24051 24091 24131 24171 24211 24251 24291 24331 24371 24411 24451 24491 24531 24571 24611 24651 24691 24731 24771 24811 24851 24891 24931 24971 25011 25051 25091 25131 25171 25211 25251 25291 25331 25371 25411 25451 25491 25531 25571 25611 25651 25691 25731 25771 25811 25851 25891 25931 25971 26011 26051 26091 26131 26171 26211 26251 26291 26331 26371 26411 26451 26491 26531 26571 26611 26651 26691 26731 26771 26811 26851 26891 26931 26971 27011 27051 27091 27131 27171 27211 27251 27291 27331 27371 27411 27451 27491 27531 27571 27611 27651 27691 27731 27771 27811 27851 27891 27931 27971 28011 28051 28091 28131 28171 28211 28251 28291 28331 28371 28411 28451 28491 28531 28571 28611 28651 28691 28731 28771 28811 28851 28891 28931 28971 29011 29051 29091 29131 29171 29211 29251 29291 29331 29371 29411 29451 29491 29531 29571 29611 29651 29691 29731 29771 29811 29851 29891 29931 29971 30011 30051 30091 30131 30171 30211 30251 30291 30331 30371 30411 30451 30491 30531 30571 30611 30651 30691 30731 30771 30811 30851 30891 30931 30971 31011 31051 31091 31131 31171 31211 31251 31291 31331 31371 31411 31451 31491 31531 31571 31611 31651 31691 31731 31771 31811 31851 31891 31931 31971 32011 32051 32091 32131 32171 32211 32251 32291 32331 32371 32411 32451 32491 32531 32571 32611 32651 32691 32731 32771 32811 32851 32891 32931 32971 33011 33051 33091 33131 33171 33211 33251 33291 33331 33371 33411 33451 33491 33531 33571 33611 33651 33691 33731 33771 33811 33851 33891 33931 33971 34011 34051 34091 34131 34171 34211 34251 34291 34331 34371 34411 34451 34491 34531 34571 34611 34651 34691 34731 34771 34811 34851 34891 34931 34971 35011 35051 35091 35131 35171 35211 35251 35291 35331 35371 35411 35451 35491 35531 35571 35611 35651 35691 35731 35771 35811 35851 35891 35931 35971 36011 36051 36091 36131 36171 36211 36251 36291 36331 36371 36411 36451 36491 36531 36571 36611 36651 36691 36731 36771 36811 36851 36891 36931 36971 37011 37051 37091 37131 37171 37211 37251 37291 37331 37371 37411 37451 37491 37531 37571 37611 37651 37691 37731 37771 37811 37851 37891 37931 37971 38011 38051 38091 38131 38171 38211 38251 38291 38331 38371 38411 38451 38491 38531 38571 38611 38651 38691 38731 38771 38811 38851 38891 38931 38971 39011 39051 39091 39131 39171 39211 39251 39291 39331 39371 39411 39451 39491 39531 39571 39611 39651 39691 39731 39771 39811 39851 39891 39931 39971 40011 40051 40091 40131 40171 40211 40251 40291 40331 40371 40411 40451 40491 40531 40571 40611 40651 40691 40731 40771 40811 40851 40891 40931 40971 41011 41051 41091 41131 41171 41211 41251 41291 41331 41371 41411 41451 41491 41531 41571 41611 41651 41691 41731 41771 41811 41851 41891 41931 41971 42011 42051 42091 42131 42171 42211 42251 42291 42331 42371 42411 42451 42491 42531 42571 42611 42651 42691 42731 42771 42811 42851 42891 42931 42971 43011 43051 43091 43131 43171 43211 43251 43291 43331 43371 43411 43451 43491 43531 43571 43611 43651 43691 43731 43771 43811 43851 43891 43931 43971 44011 44051 44091 44131 44171 44211 44251 44291 44331 44371 44411 44451 44491 44531 44571 44611 44651 44691 44731 44771 44811 44851 44891 44931 44971 45011 45051 45091 45131 45171 45211 45251 45291 45331 45371 45411 45451 45491 45531 45571 45611 45651 45691 45731 45771 45811 45851 45891 45931 45971 46011 46051 46091 46131 46171 46211 46251 46291 46331 46371 46411 46451 46491 46531 46571 46611 46651 46691 46731 46771 46811 46851 46891 46931 46971 47011 47051 47091 47131 47171 47211 47251 47291 47331 47371 47411 47451 47491 47531 47571 47611 47651 47691 47731 47771 47811 47851 47891 47931 47971 48011 48051 48091 48131 48171 48211 48251 48291 48331 48371 48411 48451 48491 48531 48571 48611 48651 48691 48731 48771 48811 48851 48891 48931 48971 49011 49051 49091 49131 49171 49211 49251 49291 49331 49371 49411 49451 49491 49531 49571 49611 49651 49691 49731 49771 49811 49851 49891 49931 49971 50011 50051 50091 50131 50171 50211 50251 50291 50331 50371 50411 50451 50491 50531 50571 50611 50651 50691 50731 50771 50811 50851 50891 50931 50971 51011 51051 51091 51131 51171 51211 51251 51291 51331 51371 51411 51451 51491 51531 51571 51611 51651 51691 51731 51771 51811 51851 51891 51931 51971 52011 52051 52091 52131 52171 52211 52251 52291 52331 52371 52411 52451 52491 52531 52571 52611 52651 52691 52731 52771 52811 52851 52891 52931 52971 53011 53051 53091 53131 53171 53211 53251 53291 53331 53371 53411 53451 53491 53531 53571 53611 53651 53691 53731 53771 53811 53851 53891 53931 53971 54011 54051 54091 54131 54171 54211 54251 54291 54331 54371 54411 54451 54491 54531 54571 54611 54651 54691 54731 54771 54811 54851 54891 54931 54971 55011 55051 55091 55131 55171 55211 55251 55291 55331 55371 55411 55451 55491 55531 55571 55611 55651 55691 55731 55771 55811 55851 55891 55931 55971 56011 56051 56091 56131 56171 56211 56251 56291 56331 56371 56411 56451 56491 56531 56571 56611 56651 56691 56731 56771 56811 56851 56891 56931 56971 57011 57051 57091 57131 57171 57211 57251 57291 57331 57371 57411 57451 57491 57531 57571 57611 57651 57691 57731 57771 57811 57851 57891 57931 57971 58011 58051 58091 58131 58171 58211 58251 58291 58331 58371 58411 58451 58491 58531 58571 58611 58651 58691 58731 58771 58811 58851 58891 58931 58971 59011 59051 59091 59131 59171 59211 59251 59291 59331 59371 59411 59451 59491 59531 59571 59611 59651 59691 59731 59771 59811 59851 59891 59931 59971 60011 60051 60091 60131 60171 60211 60251 60291 60331 60371 60411 60451 60491 60531 60571 60611 60651 60691 60731 60771 60811 60851 60891 60931 60971 61011 61051 61091 61131 61171 61211 61251 61291 61331 61371 61411 61451 61491 61531 61571 61611 61651 61691 61731 61771 61811 61851 61891 61931 61971 62011 62051 62091 62131 62171 62211 62251 62291 62331 62371 62411 62451 62491 62531 62571 62611 62651 62691 62731 62771 62811 62851 62891 62931 62971 63011 63051 63091 63131 63171 63211 63251 63291 63331 63371 63411 63451 63491 63531 63571 63611 63651 63691 63731 63771 63811 63851 63891 63931 63971 64011 64051 64091 64131 64171 64211 64251 64291 64331 64371 64411 64451 64491 64531 64571 64611 64651 64691 64731 64771 64811 64851 64891 64931 64971 65011 65051 65091 65131 65171 65211 65251 65291 65331 65371 65411 65451 65491 65531 65571 65611 65651 65691 65731 65771 65811 65851 65891 65931 65971 66011 66051 66091 66131 66171 66211 66251 66291 66331 66371 66411 66451 66491 66531 66571 66611 66651 66691 66731 66771 66811 66851 66891 66931 66971 67011 67051 67091 67131 67171 67211 67251 67291 67331 67371 67411 67451 67491 67531 67571 67611 67651 67691 67731 67771 67811 67851 67891 67931 67971 68011 68051 68091 68131 68171 68211 68251 68291 68331 68371 68411 68451 68491 68531 68571 68611 68651 68691 68731 68771 68811 68851 68891 68931 68971 69011 69051 69091 69131 69171 69211 69251 69291 69331 69371 69411 69451 69491 69531 69571 69611 69651 69691 69731 69771 69811 69851 69891 69931 69971 70011 70051 70091 70131 70171 70211 70251 70291 70331 70371 70411 70451 70491 70531 70571 70611 70651 70691 70731 70771 70811 70851 70891 70931 70971 71011 71051 71091 71131 71171 71211 71251 71291 71331 71371 71411 71451 71491 71531 71571 71611 71651 71691 71731 71771 71811 71851 71891 71931 71971 72011 72051 72091 72131 72171 72211 72251 72291 72331 72371 72411 72451 72491 72531 72571 72611 72651 72691 72731 72771 72811 72851 72891 72931 72971 73011 73051 73091 73131 73171 73211 73251 73291 73331 73371 73411 73451 73491 73531 73571 73611 73651 73691 73

| FT | DISULFID | 849  | 859  | BY SIMILARITY. |
|----|----------|------|------|----------------|
| FT | DISULFID | 854  | 868  | BY SIMILARITY. |
| FT | DISULFID | 870  | 883  | BY SIMILARITY. |
| FT | DISULFID | 952  | 964  | BY SIMILARITY. |
| FT | DISULFID | 959  | 973  | BY SIMILARITY. |
| FT | DISULFID | 975  | 988  | BY SIMILARITY. |
| FT | DISULFID | 1070 | 1082 | BY SIMILARITY. |
| FT | DISULFID | 1077 | 1091 | BY SIMILARITY. |
| FT | DISULFID | 1093 | 1106 | BY SIMILARITY. |
| FT | DISULFID | 1112 | 1124 | BY SIMILARITY. |
| FT | DISULFID | 1119 | 1133 | BY SIMILARITY. |
| FT | DISULFID | 1135 | 1149 | BY SIMILARITY. |
| FT | DISULFID | 1155 | 1167 | BY SIMILARITY. |
| FT | DISULFID | 1162 | 1176 | BY SIMILARITY. |
| FT | DISULFID | 1178 | 1191 | BY SIMILARITY. |
| FT | DISULFID | 1197 | 1209 | BY SIMILARITY. |
| FT | DISULFID | 1204 | 1218 | BY SIMILARITY. |
| FT | DISULFID | 1220 | 1233 | BY SIMILARITY. |
| FT | DISULFID | 1239 | 1250 | BY SIMILARITY. |
| FT | DISULFID | 1246 | 1259 | BY SIMILARITY. |
| FT | DISULFID | 1261 | 1274 | BY SIMILARITY. |
| FT | DISULFID | 1280 | 1292 | BY SIMILARITY. |
| FT | DISULFID | 1287 | 1301 | BY SIMILARITY. |
| FT | DISULFID | 1303 | 1316 | BY SIMILARITY. |
| FT | DISULFID | 1322 | 1334 | BY SIMILARITY. |
| FT | DISULFID | 1329 | 1343 | BY SIMILARITY. |
| FT | DISULFID | 1345 | 1358 | BY SIMILARITY. |
| FT | DISULFID | 1364 | 1377 | BY SIMILARITY. |
| FT | DISULFID | 1371 | 1386 | BY SIMILARITY. |
| FT | DISULFID | 1388 | 1399 | BY SIMILARITY. |
| FT | DISULFID | 1405 | 1418 | BY SIMILARITY. |
| FT | DISULFID | 1412 | 1427 | BY SIMILARITY. |
| FT | DISULFID | 1429 | 1440 | BY SIMILARITY. |
| FT | DISULFID | 1446 | 1458 | BY SIMILARITY. |
| FT | DISULFID | 1453 | 1467 | BY SIMILARITY. |
| FT | DISULFID | 1469 | 1482 | BY SIMILARITY. |
| FT | DISULFID | 1488 | 1499 | BY SIMILARITY. |
| FT | DISULFID | 1494 | 1508 | BY SIMILARITY. |
| FT | DISULFID | 1510 | 1523 | BY SIMILARITY. |
| FT | DISULFID | 1529 | 1540 | BY SIMILARITY. |
| FT | DISULFID | 1535 | 1549 | BY SIMILARITY. |
| FT | DISULFID | 1551 | 1564 | BY SIMILARITY. |
| FT | DISULFID | 1647 | 1659 | BY SIMILARITY. |
| FT | DISULFID | 1654 | 1668 | BY SIMILARITY. |
| FT | DISULFID | 1670 | 1683 | BY SIMILARITY. |
| FT | DISULFID | 1689 | 1701 | BY SIMILARITY. |

Query Match 13.2%; Score 373; DB 1; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 4.3e-17;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

|    |      |   |                          |
|----|------|---|--------------------------|
| Qy | 226  | DCSYENGCGCHACNAIPGAPRCQCPAGAAQQADGRSGCTASATQSCNDLCEHFCVPNDP | -- 283                   |
| Db | 1238 | ECMTMGNGCDTQCTNSBSGYECSSEGYALMPGRSCA                        | -----DIDE--CENNPDIC 1287 |
| Qy | 284  | -----QPSYSYCMCBTGYRLAADOHRCBBDVDDCILEPSPCP-QRCVNTQGGPCHCYP  | 335                      |
| Db | 1288 | DGGQCTNIPGEYRCLCYDGFMA5MDMKTCTDVNECDLNPICMGECENTKGSFCHCQL   | 1347                     |
| Qy | 336  | NYDLVDGE--CVBPVDPDC--FRANCEYQCOPLN-QTSYLVCVABGFA            | -----PIP 380             |
| Db | 1348 | GY5VKKGTTGCTD-VBECEYGAHNCMDHMASCLNVPGSFKSCREGWVGKICIDLDECA  | 1406                     |
| Qy | 381  | HEPHRCOMFCNQACPADCDPNTQAS--CECPGYILDDGFICTDIDE              | -----CBNG- 430           |
| Db | 1407 | NGTHQCSI-----NAQC-VNTPGSYRACSSGF-TGDGFTCSVDVDECAENTNLCENGQ  | 1457                     |
| Qy | 431  | -----GF-----CS-----GVCHNLPGTFFCICGPD5ALV                    | 455                      |
| Db | 1458 | CLNVPAYRCEBNGFTPASDRSCQDIDCSFQNICVFTGTCNNLPGMFHICDDGYELD    | 1517                     |
| Qy | 456  | RHTG--TDCC 463  |                          |

DR PROSITE; PS00022; EGF 1; FALSE\_NEG.  
 DR PROSITE; PS01386; EGF 2; 5.  
 DR PROSITE; PS00026; EGF 3; 4.  
 DR PROSITE; PS01187; EGF CA; 9.  
 KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
 KW Calcium-binding; Repeat.  
 FT SIGNAL 1 27  
 FT CHAIN 28 1184  
 FT DOMAIN 28 444  
 FT DOMAIN 28 444  
 FT DOMAIN 178 177  
 FT DOMAIN 178 444  
 FT DOMAIN 448 480  
 FT DOMAIN 448 519  
 FT DOMAIN 521 553  
 FT DOMAIN 604 645  
 FT DOMAIN 679 718  
 FT DOMAIN 719 763  
 FT DOMAIN 764 809  
 FT DOMAIN 810 857  
 FT DOMAIN 858 900  
 FT DOMAIN 901 942  
 FT DOMAIN 943 981  
 FT DOMAIN 982 1024  
 FT DOMAIN 1025 1069  
 FT DOMAIN 1070 1184  
 FT DISULFID 445 472  
 FT DISULFID 446 479  
 FT DISULFID 459 480  
 FT DISULFID 489 518  
 FT DISULFID 502 519  
 FT DISULFID 521 545  
 FT DISULFID 522 552  
 FT DISULFID 535 553  
 FT DISULFID 608 620  
 FT DISULFID 616 629  
 FT DISULFID 631 644  
 FT DISULFID 683 693  
 FT DISULFID 689 702  
 FT DISULFID 704 717  
 FT DISULFID 723 736  
 FT DISULFID 730 745  
 FT DISULFID 751 762  
 FT DISULFID 768 781  
 FT DISULFID 775 790  
 FT DISULFID 796 808  
 FT DISULFID 814 827  
 FT DISULFID 821 836  
 FT DISULFID 843 856  
 FT DISULFID 862 875  
 FT DISULFID 869 884  
 FT DISULFID 886 899  
 FT DISULFID 905 917  
 FT DISULFID 913 926  
 FT DISULFID 928 941  
 FT DISULFID 947 956  
 FT DISULFID 952 965  
 FT DISULFID 967 980  
 FT DISULFID 986 998  
 FT DISULFID 994 1007  
 FT DISULFID 1009 1023  
 FT DISULFID 1029 1042  
 FT DISULFID 1036 1051  
 FT DISULFID 1056 1068  
 FT CARBOHYD 180 180  
 FT CARBOHYD 507 507  
 FT CARBOHYD 1035 1035  
 FT SEQUENCE 1184 AA; 126543 MW; CA48490A5F9EC5D CRC64;  
 Query Match 13.0%; Score 367.5; DB 1; Length 1184;  
 Best Local Similarity 30.0%; Pred. No. 4.1e-17;  
 Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
 QY 136 EEQQCEVKAD-GFLCEHFPATC-----RPLAV-----EPGAAAASVITYGTFPAARG 183

Db 531 EQQCESNPNGYPCN-HVMLSCCEGEEPLIPEVRRPPEPAAAPRRVS-----EAM 582  
 QY 184 ADFQALPVGSSAAV---APLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCGCHACNA 240  
 Db 583 AGREALSLGTEALPNSLPFGDDQDECLLLPGEL-----CQHLGIN 622  
 QY 241 IPGAPROCPCAGALQADGRSC-----TASAT-----QS 269  
 Db 623 TVGSYHCACFFGFSQLDDGRTCPGHPQPPEAQEPALKSEFSEFVAVNTIPLPQNT 682  
 QY 270 CND--LCEHFCVPNDQPGSYSCMCTGYRLAAQHRCEVDVDCILEPSPCP--QRCVNT 325  
 Db 683 CKDNGPKQVC---STVGSSAICSCFFGYALMADGVSCEDINECVTDLHTCSRGEHCVT 739  
 QY 326 QGSEFCH---CYPNYDLVDCGCVPEVDPDPCFRANCEYOCQPLNQTSLVCV----- 371  
 Db 740 LGSFHYKALTCERGYALKQEC-EDVDEC--AMGTHTCQP-----GFLQNTKGSFYCQA 792  
 QY 372 ---CAEGFAPIPH-----EPHRCQMFNCNOTACPADCDPNTQASCEGPEV-I 414  
 Db 793 RQRCMDGFLQDPGEGNCVDINECTSLSPCRPGFSCINTVGSYTCORNPLI---CARGTHA 849  
 QY 415 LDGFICTDIDECENGQFCSG---VCHNLPGETPECIC 448  
 Db 850 SDGAKCDVNECETGVHRCGEGQVCHNLPFSYRDC 886

RESULT 9  
 FBNI\_BOVIN  
 ID FBNI\_BOVIN STANDARD; PRT; 2871 AA.  
 AC P98133;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor (MF340).  
 GN FBNI.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=95137597; PubMed=7835900;  
 RA Tilstra D.J., Potter K.A., Byers P.H.;  
 RT "Sequence of the coding region of the bovine fibrillin cDNA and  
 RT localization to bovine chromosome 10.";  
 RL Genomics 23:480-485(1994).  
 RN [2]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=96132851; PubMed=8557636;  
 RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,  
 RA Nicholl J.K., Sutherland G.R., Cleary E.G.;  
 RT "Further characterization of proteins associated with elastic fiber  
 RT microfibrils including the molecular cloning of MAGP-2 (MP25).";  
 RL J. Biol. Chem. 271:1096-1103(1996).  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 CC that binds calcium. Fibrillin-1-containing microfibrils provide  
 CC long-term force bearing structural support.  
 CC -!- PTM: Forms intermolecular disulfide bonds either with other  
 CC fibrillin-1 molecules or with other components of the  
 CC microfibrils.  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBRP) domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch).

-----

EMBL; L28748; AAA74122.1; -

PIR; A55567; A55567.

HSSP; P35555; IAPJ.

InterPro; IPR000152; Asx hydroxyl\_s.

InterPro; IPR001881; EGF\_Ca.

InterPro; IPR006205; EGF\_like.

InterPro; IPR002212; Fibril-assoc.

Pfam; PF00008; EGF; 46.

Pfam; PF00683; TB; 9.

SMART; SM00179; EGF\_CA; 42.

PROSITE; PS00010; ASX HYDROXYL; 43.

PROSITE; PS00022; EGF\_1; 2.

PROSITE; PS01186; EGF\_2; 38.

PROSITE; PS00026; EGF\_3; 45.

PROSITE; PS01187; EGF\_CA; 43.

KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;

KW Repeat; Signal; Multigene family;

FT SIGNAL 1 27 POTENTIAL.

FT CHAIN 28 2871 FIBRILLIN 1.

FT DOMAIN 81 112 EGF-LIKE 1.

FT DOMAIN 115 146 EGF-LIKE 2.

FT DOMAIN 147 178 EGF-LIKE 3.

FT DOMAIN 246 287 EGF-LIKE 4.

FT DOMAIN 288 329 EGF-LIKE 5.

FT REPEAT 330 390 TGFBP 1.

FT DOMAIN 392 446 PRO-RICH.

FT DOMAIN 449 489 EGF-LIKE 6.

FT DOMAIN 490 529 EGF-LIKE 7.

FT DOMAIN 530 571 EGF-LIKE 8.

FT DOMAIN 572 612 EGF-LIKE 9.

FT DOMAIN 613 653 EGF-LIKE 10.

FT REPEAT 654 722 TGFBP 2.

FT DOMAIN 723 764 EGF-LIKE 11.

FT DOMAIN 765 806 EGF-LIKE 12.

FT DOMAIN 807 846 EGF-LIKE 13.

FT DOMAIN 910 951 EGF-LIKE 14.

FT REPEAT 952 1027 TGFBP 3.

FT DOMAIN 1028 1069 EGF-LIKE 15.

FT DOMAIN 1070 1112 EGF-LIKE 16.

FT DOMAIN 1113 1154 EGF-LIKE 17.

FT DOMAIN 1155 1196 EGF-LIKE 18.

FT DOMAIN 1197 1237 EGF-LIKE 19.

FT DOMAIN 1238 1279 EGF-LIKE 20.

FT DOMAIN 1280 1321 EGF-LIKE 21.

FT DOMAIN 1322 1362 EGF-LIKE 22.

FT DOMAIN 1363 1403 EGF-LIKE 23.

FT DOMAIN 1404 1445 EGF-LIKE 24.

FT DOMAIN 1446 1486 EGF-LIKE 25.

FT DOMAIN 1487 1527 EGF-LIKE 26.

FT REPEAT 1528 1605 TGFBP 4.

FT DOMAIN 1606 1647 EGF-LIKE 27.

FT DOMAIN 1648 1688 EGF-LIKE 28.

FT REPEAT 1689 1765 TGFBP 5.

FT DOMAIN 1766 1807 EGF-LIKE 29.

FT DOMAIN 1808 1848 EGF-LIKE 30.

FT DOMAIN 1849 1890 EGF-LIKE 31.

FT DOMAIN 1891 1929 EGF-LIKE 32.

FT DOMAIN 1930 1972 EGF-LIKE 33.

FT DOMAIN 1973 2012 EGF-LIKE 34.

FT DOMAIN 2013 2054 EGF-LIKE 35.

FT REPEAT 2055 2126 TGFBP 6.

FT DOMAIN 2127 2165 EGF-LIKE 36.

FT DOMAIN 2166 2205 EGF-LIKE 37.

FT DOMAIN 2206 2246 EGF-LIKE 38.

FT DOMAIN 2247 2290 EGF-LIKE 39.

FT DOMAIN 2291 2332 EGF-LIKE 40.

FT REPEAT 2333 2401 TGFBP 7.

FT DOMAIN 2402 2443 EGF-LIKE 41.

FT DOMAIN 2444 2484 EGF-LIKE 42.

FT DOMAIN 2485 2523 EGF-LIKE 43.

FT DOMAIN 2524 2566 EGF-LIKE 44.

FT DOMAIN 2567 2607 EGF-LIKE 45.

FT DOMAIN 2607 2647 EGF-LIKE 46.

FT DOMAIN 2648 2687 EGF-LIKE 47.

FT DISULFID 85 94 BY SIMILARITY.

FT DISULFID 89 100 BY SIMILARITY.

FT DISULFID 102 111 BY SIMILARITY.

FT DISULFID 119 129 BY SIMILARITY.

FT DISULFID 123 134 BY SIMILARITY.

FT DISULFID 136 145 BY SIMILARITY.

FT DISULFID 150 160 BY SIMILARITY.

FT DISULFID 154 166 BY SIMILARITY.

FT DISULFID 168 177 BY SIMILARITY.

FT DISULFID 250 262 BY SIMILARITY.

FT DISULFID 257 271 BY SIMILARITY.

FT DISULFID 273 286 BY SIMILARITY.

FT DISULFID 292 304 BY SIMILARITY.

FT DISULFID 299 313 BY SIMILARITY.

FT DISULFID 315 328 BY SIMILARITY.

FT DISULFID 433 465 BY SIMILARITY.

FT DISULFID 460 474 BY SIMILARITY.

FT DISULFID 476 488 BY SIMILARITY.

FT DISULFID 494 504 BY SIMILARITY.

FT DISULFID 499 513 BY SIMILARITY.

FT DISULFID 515 528 BY SIMILARITY.

FT DISULFID 534 546 BY SIMILARITY.

FT DISULFID 541 555 BY SIMILARITY.

FT DISULFID 557 570 BY SIMILARITY.

FT DISULFID 576 587 BY SIMILARITY.

FT DISULFID 582 596 BY SIMILARITY.

FT DISULFID 598 611 BY SIMILARITY.

FT DISULFID 617 628 BY SIMILARITY.

FT DISULFID 623 637 BY SIMILARITY.

FT DISULFID 639 652 BY SIMILARITY.

FT DISULFID 727 739 BY SIMILARITY.

FT DISULFID 734 748 BY SIMILARITY.

FT DISULFID 750 763 BY SIMILARITY.

FT DISULFID 769 781 BY SIMILARITY.

FT DISULFID 776 790 BY SIMILARITY.

FT DISULFID 792 805 BY SIMILARITY.

FT DISULFID 811 821 BY SIMILARITY.

FT DISULFID 816 830 BY SIMILARITY.

FT DISULFID 832 845 BY SIMILARITY.

FT DISULFID 914 926 BY SIMILARITY.

FT DISULFID 921 935 BY SIMILARITY.

FT DISULFID 937 950 BY SIMILARITY.

FT DISULFID 1032 1044 BY SIMILARITY.

FT DISULFID 1039 1053 BY SIMILARITY.

FT DISULFID 1055 1068 BY SIMILARITY.

FT DISULFID 1074 1085 BY SIMILARITY.

FT DISULFID 1081 1095 BY SIMILARITY.

FT DISULFID 1097 1111 BY SIMILARITY.

FT DISULFID 1117 1129 BY SIMILARITY.

FT DISULFID 1124 1138 BY SIMILARITY.

FT DISULFID 1140 1153 BY SIMILARITY.

FT DISULFID 1159 1171 BY SIMILARITY.

FT DISULFID 1166 1180 BY SIMILARITY.

FT DISULFID 1182 1195 BY SIMILARITY.

FT DISULFID 1201 1212 BY SIMILARITY.

FT DISULFID 1208 1221 BY SIMILARITY.

FT DISULFID 1223 1236 BY SIMILARITY.

FT DISULFID 1242 1254 BY SIMILARITY.

FT DISULFID 1249 1263 BY SIMILARITY.

FT DISULFID 1265 1278 BY SIMILARITY.

FT DISULFID 1284 1296 BY SIMILARITY.

FT DISULFID 1291 1305 BY SIMILARITY.

FT DISULFID 1307 1320 BY SIMILARITY.

FT DISULFID 1326 1339 BY SIMILARITY.

FT DISULFID 1333 1348 BY SIMILARITY.

FT DISULFID 1350 1361 BY SIMILARITY.

FT DISULFID 1367 1380 BY SIMILARITY.

FT DISULFID 1374 1389 BY SIMILARITY.

FT DISULFID 1391 1402 BY SIMILARITY.

FT DISULFID 1408 1420 BY SIMILARITY.



```
FT DOMAIN 2402 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 94 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 111 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 597 BY SIMILARITY.
FT DISULFID 592 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.

Query Match 12.9%; Score 366; DB 1; Length 2871;
Best Local Similarity 30.6%; Pred. No. 1.2e-16;
Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

QY 226 DCSVENGGCHACNAIPGAFRCQCPAGAAQADGRSTASATQSCNDLCEHFCVNP--- 282
DB 1200 ECSIMNGGCEFTCTNSEGSEYECSCQPGFALPMDQORCT-----DIDE--CEDPNIC 1249
QY 283 -----DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSC--PORCVNTQGGFECHCYP 335
DB 1250 DGGQCTNIPGEYRCLCYDGFWASEDMKTCVDVNECDLNPICLSGTCENTKSGFICHCDM 1309
QY 336 NYDLVDGE--CPEPVDFC--FRANCEYQCOFLNQT--SYLCVCAEGFA-----PIP 380
DB 1310 GYSGKKGKGTCTD--INECEIGAHCNDRHAVCTNTAGSFNCSCSPGWIQDKICTDDECS 1368
QY 381 HEHRCQMFQNTACPADCPNTQAS--CECPGYILDGFICTDIDEC-EN----- 429
DB 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCADLDECSNVKLCGNVQ 1419
QY 430 -----GGF-----CS-----GVCHNLPGTFECICGPDALV 455
DB 1420 CLYAPGGYHCEYDMGFVPSADRSKSCVDSDECSLPNCTCVFTCHNLPGLFRCEC----- 1472
QY 456 RHGTGDCDSKVDGSGSGGSEPPPS-----PTPGS-----TLTPPAVGLV 495
DB 1473 -EIGYELDRSGGNCCTDVNECLEPPTCISGNVCNVTGSYTCVCPDFELNPTRVGCV 1527

RESULT 11
FBN2 HUMAN
ID FBN2 HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apffelroth S.D., Hu W., Davis E.C., Sanguineti C.,
RA Bonadio J., Mecham R.P., Ramirez F.
RT "Structure and expression of fibrillin-2, a novel microfibrillar
```



RT component preferentially located in elastic matrices.";

RL J. Cell Biol. 124:855-863(1994).

RN [2]

RP SEQUENCE OF 752-1505 FROM N.A.

RX MEDLINE=91304567; PubMed=1852206;

RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,

RT Tsipouras P., Ramirez F., Hollister D.;

RT "Linkage of Marfan syndrome and a phenotypically related disorder to

RT two different fibrillin genes.";

RL Nature 352:330-334(1991).

RN [3]

RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.

RX MEDLINE=96083599; PubMed=7493032;

RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;

RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,

RT congenital contractual arachnodactyly.";

RL Nat. Genet. 11:456-458(1995).

RN [4]

RP VARIANTS CCA HIS-1114.

RX MEDLINE=98407789; PubMed=9737771;

RA Babcock D., Gasner C., Francke U., Maalen C.;

RT "A single mutation that results in an asp-to-his substitution and

RT partial exon skipping in a family with congenital contractual

RT arachnodactyly.";

RL Hum. Genet. 103:22-28(1998).

RN [5]

RP VARIANTS CCA PHE-1141 AND TRP-1252.

RX MEDLINE=20259236; PubMed=10797416;

RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,

RT Godfrey M.;

RT "Two novel fibrillin-2 mutations in congenital contractual

RT arachnodactyly.";

RL Am. J. Med. Genet. 92:7-12(2000).

CC -1- FUNCTION: Structural component of connective tissue microfibrils

CC that binds calcium. Fibrillin-2-containing microfibrils regulate

CC the early process of elastic fiber assembly.

CC -1- DISEASE: Defects in FBN2 are the cause of congenital contractual

CC arachnodactyly (CCA) [MIM:121050]; also known as Beals syndrome.

CC CCA is phenotypically similar to Marfan syndrome, but does not

CC effect the aorta and the eyes.

CC -1- SIMILARITY: Contains 47 EGF-like domains.

CC -1- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.

CC -1- DATABASE: NAME=Elastic Fiber Homepage; NOTE=Fibrillin 2 page;

CC WWW="http://ef.wustl.edu/genes/FBN2.htm".

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; U03272; AAA18950.1; -.

DR EMBL; X62009; -; NOT\_ANNOTATED\_CDS.

DR PIR; A54105; A54105.

DR HSSP; P35555; 1EMN.

DR Genew; HGNC:3604; FBN2.

DR MIM; 121050; -.

DR GO; GO:0005578; C:extracellular matrix; TAS.

DR GO; GO:0005201; F:extracellular matrix structural constituent; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.

DR InterPro; IPR000152; ASX\_hydroxyl\_S.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR001438; EGF\_II.

DR InterPro; IPR006209; EGF\_III.

DR InterPro; IPR002212; Fibril-assoc.

DR Pfam; PF000683; TB\_9.

DR PRINTS; PR00010; EGFBL00D.

DR SMART; SM00179; EGF\_CA; 43.

DR PROSITE; PS00010; ASX\_HYDROXYL; 43.

DR PROSITE; PS00022; EGF\_1; 2.

DR PROSITE; PS01186; EGF\_2; 37.

DR PROSITE; PS00026; EGF\_3; 45.

DR PROSITE; PS01187; EGF\_CA; 42.

KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;

KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.

FT SIGNAL 1 28

FT CHAIN 29 2911

FT DOMAIN 111 142

FT EGF-LIKE 1.

FT EGF-LIKE 2.

FT EGF-LIKE 3.

FT EGF-LIKE 4.

FT EGF-LIKE 5.

FT EGF-LIKE 6.

FT EGF-LIKE 7.

FT EGF-LIKE 8.

FT EGF-LIKE 9.

FT EGF-LIKE 10.

FT EGF-LIKE 11.

FT EGF-LIKE 12.

FT EGF-LIKE 13.

FT EGF-LIKE 14.

FT EGF-LIKE 15.

FT EGF-LIKE 16.

FT EGF-LIKE 17.

FT EGF-LIKE 18.

FT EGF-LIKE 19.

FT EGF-LIKE 20.

FT EGF-LIKE 21.

FT EGF-LIKE 22.

FT EGF-LIKE 23.

FT EGF-LIKE 24.

FT EGF-LIKE 25.

FT EGF-LIKE 26.

FT EGF-LIKE 27.

FT EGF-LIKE 28.

FT EGF-LIKE 29.

FT EGF-LIKE 30.

FT EGF-LIKE 31.

FT EGF-LIKE 32.

FT EGF-LIKE 33.

FT EGF-LIKE 34.

FT EGF-LIKE 35.

FT EGF-LIKE 36.

FT EGF-LIKE 37.

FT EGF-LIKE 38.

FT EGF-LIKE 39.

FT EGF-LIKE 40.

FT EGF-LIKE 41.

FT EGF-LIKE 42.

FT EGF-LIKE 43.

FT EGF-LIKE 44.

FT EGF-LIKE 45.

FT EGF-LIKE 46.

FT EGF-LIKE 47.

FT EGF-LIKE 48.

FT EGF-LIKE 49.

FT EGF-LIKE 50.

FT EGF-LIKE 51.

FT EGF-LIKE 52.

FT EGF-LIKE 53.

FT EGF-LIKE 54.

FT EGF-LIKE 55.

FT EGF-LIKE 56.

FT EGF-LIKE 57.

FT EGF-LIKE 58.

FT EGF-LIKE 59.

FT EGF-LIKE 60.

FT EGF-LIKE 61.

FT EGF-LIKE 62.

FT EGF-LIKE 63.

FT EGF-LIKE 64.

FT EGF-LIKE 65.

FT EGF-LIKE 66.

FT EGF-LIKE 67.

FT EGF-LIKE 68.

FT EGF-LIKE 69.

FT EGF-LIKE 70.

FT EGF-LIKE 71.

FT EGF-LIKE 72.

FT EGF-LIKE 73.

FT EGF-LIKE 74.

FT EGF-LIKE 75.

FT EGF-LIKE 76.

FT EGF-LIKE 77.

FT EGF-LIKE 78.

FT EGF-LIKE 79.

FT EGF-LIKE 80.

FT EGF-LIKE 81.

FT EGF-LIKE 82.

FT EGF-LIKE 83.

FT EGF-LIKE 84.

FT EGF-LIKE 85.

FT EGF-LIKE 86.

FT EGF-LIKE 87.

FT EGF-LIKE 88.

FT EGF-LIKE 89.

FT EGF-LIKE 90.

FT EGF-LIKE 91.

FT EGF-LIKE 92.

FT EGF-LIKE 93.

FT EGF-LIKE 94.

FT EGF-LIKE 95.

FT EGF-LIKE 96.

FT EGF-LIKE 97.

FT EGF-LIKE 98.

FT EGF-LIKE 99.

FT EGF-LIKE 100.

FT EGF-LIKE 101.

FT EGF-LIKE 102.

FT EGF-LIKE 103.

FT EGF-LIKE 104.

FT EGF-LIKE 105.

FT EGF-LIKE 106.

FT EGF-LIKE 107.

FT EGF-LIKE 108.

FT EGF-LIKE 109.

FT EGF-LIKE 110.

FT EGF-LIKE 111.

FT EGF-LIKE 112.

FT EGF-LIKE 113.

FT EGF-LIKE 114.

FT EGF-LIKE 115.

FT EGF-LIKE 116.

FT EGF-LIKE 117.

FT EGF-LIKE 118.

FT EGF-LIKE 119.

FT EGF-LIKE 120.

FT EGF-LIKE 121.

FT EGF-LIKE 122.

FT EGF-LIKE 123.

FT EGF-LIKE 124.

FT EGF-LIKE 125.

FT EGF-LIKE 126.

FT EGF-LIKE 127.

FT EGF-LIKE 128.

FT EGF-LIKE 129.

FT EGF-LIKE 130.

FT EGF-LIKE 131.

FT EGF-LIKE 132.

FT EGF-LIKE 133.

FT EGF-LIKE 134.

FT EGF-LIKE 135.

FT EGF-LIKE 136.

FT EGF-LIKE 137.

FT EGF-LIKE 138.

FT EGF-LIKE 139.

FT EGF-LIKE 140.

FT EGF-LIKE 141.

FT EGF-LIKE 142.

FT EGF-LIKE 143.

FT EGF-LIKE 144.

FT EGF-LIKE 145.

FT EGF-LIKE 146.

FT EGF-LIKE 147.

FT EGF-LIKE 148.

FT EGF-LIKE 149.

FT EGF-LIKE 150.

FT EGF-LIKE 151.

FT EGF-LIKE 152.

FT EGF-LIKE 153.

FT EGF-LIKE 154.

FT EGF-LIKE 155.

FT EGF-LIKE 156.

FT EGF-LIKE 157.

FT EGF-LIKE 158.

FT EGF-LIKE 159.

FT EGF-LIKE 160.

FT EGF-LIKE 161.

FT EGF-LIKE 162.

FT EGF-LIKE 163.

FT EGF-LIKE 164.

FT EGF-LIKE 165.

FT EGF-LIKE 166.

FT EGF-LIKE 167.

FT EGF-LIKE 168.

FT EGF-LIKE 169.

FT EGF-LIKE 170.

FT EGF-LIKE 171.

FT EGF-LIKE 172.

FT EGF-LIKE 173.

FT EGF-LIKE 174.

FT EGF-LIKE 175.

FT EGF-LIKE 176.

FT EGF-LIKE 177.

FT EGF-LIKE 178.

FT EGF-LIKE 179.

FT EGF-LIKE 180.

FT EGF-LIKE 181.

FT EGF-LIKE 182.

FT EGF-LIKE 183.

FT EGF-LIKE 184.

FT EGF-LIKE 185.

FT EGF-LIKE 186.

FT EGF-LIKE 187.

FT EGF-LIKE 188.

FT EGF-LIKE 189.

FT EGF-LIKE 190.

FT EGF-LIKE 191.

FT EGF-LIKE 192.

FT EGF-LIKE 193.

FT EGF-LIKE 194.

FT EGF-LIKE 195.

FT EGF-LIKE 196.

FT EGF-LIKE 197.

FT EGF-LIKE 198.

FT EGF-LIKE 199.

FT EGF-LIKE 200.

FT EGF-LIKE 201.

FT EGF-LIKE 202.

FT EGF-LIKE 203.

FT EGF-LIKE 204.

FT EGF-LIKE 205.

FT EGF-LIKE 206.

FT EGF-LIKE 207.

FT EGF-LIKE 208.

FT EGF-LIKE 209.

FT EGF-LIKE 210.

FT EGF-LIKE 211.

FT EGF-LIKE 212.

FT EGF-LIKE 213.

FT EGF-LIKE 214.

FT EGF-LIKE 215.

FT EGF-LIKE 216.

FT EGF-LIKE 217.

FT EGF-LIKE 218.

FT EGF-LIKE 219.

FT EGF-LIKE 220.

FT EGF-LIKE 221.

FT EGF-LIKE 222.

FT EGF-LIKE 223.

FT EGF-LIKE 224.

FT EGF-LIKE 225.

FT EGF-LIKE 226.

FT EGF-LIKE 227.

FT EGF-LIKE 228.

FT EGF-LIKE 229.

FT EGF-LIKE 230.

FT EGF-LIKE 231.

FT EGF-LIKE 232.

FT EGF-LIKE 233.

FT EGF-LIKE 234.

FT EGF-LIKE 235.

FT EGF-LIKE 236.

FT EGF-LIKE 237.

FT EGF-LIKE 238.

FT EGF-LIKE 239.

FT EGF-LIKE 240.

FT EGF-LIKE 241.

FT EGF-LIKE 242.

FT EGF-LIKE 243.

FT EGF-LIKE 244.

FT EGF-LIKE 245.

FT EGF-LIKE 246.

FT EGF-LIKE 247.

FT EGF-LIKE 248.

FT EGF-LIKE 249.

FT EGF-LIKE 250.

FT EGF-LIKE 251.

FT EGF-LIKE 252.

FT EGF-LIKE 253.

FT EGF-LIKE 254.

FT EGF-LIKE 255.

FT EGF-LIKE 256.

FT EGF-LIKE 257.

FT EGF-LIKE 258.

FT EGF-LIKE 259.

FT EGF-LIKE 260.

FT EGF-LIKE 261.

FT EGF-LIKE 262.

FT EGF-LIKE 263.

FT EGF-LIKE 264.

FT EGF-LIKE 265.

FT EGF-LIKE 266.

FT EGF-LIKE 267.

FT EGF-LIKE 268.

FT EGF-LIKE 269.

FT EGF-LIKE 270.

FT EGF-LIKE 271.

FT EGF-LIKE 272.

FT EGF-LIKE 273.

FT EGF-LIKE 274.

FT EGF-LIKE 275.

FT EGF-LIKE 276.

FT EGF-LIKE 277.

FT EGF-LIKE 278.

FT EGF-LIKE 279.

FT EGF-LIKE 280.

FT EGF-LIKE 281.

FT EGF-LIKE 282.

FT EGF-LIKE 283.

FT EGF-LIKE 284.

FT EGF-LIKE 285.

FT EGF-LIKE 286.

FT EGF-LIKE 287.

FT EGF-LIKE 288.

FT EGF-LIKE 289.

FT EGF-LIKE 290.

FT EGF-LIKE 291.

FT EGF-LIKE 292.

FT EGF-LIKE 293.

FT EGF-LIKE 294.

FT EGF-LIKE 295.

FT EGF-LIKE 296.

FT EGF-LIKE 297.

FT EGF-LIKE 298.

FT EGF-LIKE 299.

FT EGF-LIKE 300.





"Software and database for the analysis of mutations in the human FBN1 gene.",  
 RT Nucleic Acids Res. 24:137-141(1996).  
 RN [10]  
 RP REVIEW ON MFS VARIANTS.  
 RX MEDLINE=97169383; PubMed=9016526;  
 RA Collod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,  
 RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,  
 RA Richards R.I., Wang W., Junien C., Boileau C.;  
 RT "Marfan Database (second edition): software and database for the  
 RT analysis of mutations in the human FBN1 gene.",  
 RN Nucleic Acids Res. 25:147-150(1997).  
 RN [11]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=98062175; PubMed=9401003;  
 RA Hayward C., Brock D.J.H.;  
 RT "Fibrillin-1 mutations in Marfan syndrome and other type-1  
 RT fibrillinopathies.",  
 RN Hum. Mutat. 10:415-423(1997).  
 RN [12]  
 RP VARIANT MFS PRO-1137.  
 RX MEDLINE=91304569; PubMed=1852208;  
 RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,  
 RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,  
 RA Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;  
 RT "Marfan syndrome caused by a recurrent de novo missense mutation in  
 RT the fibrillin gene.",  
 RN Nature 352:337-339(1991).  
 RN [13]  
 RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.  
 RX MEDLINE=93250834; PubMed=1301946;  
 RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;  
 RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome  
 RT patients at cysteine residues in EGF-like domains.",  
 RN Hum. Mutat. 1:366-374(1992).  
 RN [14]  
 RP VARIANT MFS SER-2307.  
 RX MEDLINE=92235290; PubMed=1569206;  
 RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,  
 RA Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;  
 RT "Marfan phenotype variability in a family segregating a missense  
 RT mutation in the epidermal growth factor-like motif of the fibrillin  
 RT gene.",  
 RN J. Clin. Invest. 89:1674-1680(1992).  
 RN [15]  
 RP VARIANTS MFS ILR-548 AND ALA-723.  
 RX MEDLINE=94010946; PubMed=8406497;  
 RA Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,  
 RA Pyeritz R.E., Francomano C.A.;  
 RT "Four novel FBN1 mutations: significance for mutant transcript level  
 RT and EGF-like domain calcium binding in the pathogenesis of Marfan  
 RT syndrome.",  
 RN Genomics 17:468-475(1993).  
 RN [16]  
 RP VARIANT MFS SER-2144.  
 RX MEDLINE=93278402; PubMed=8504310;  
 RA Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;  
 RT "A novel fibrillin mutation in the Marfan syndrome which could  
 RT disrupt calcium binding of the epidermal growth factor-like module.",  
 RN Hum. Mol. Genet. 2:475-477(1993).  
 RN [17]  
 RP VARIANTS MFS ARG-862; TYR-1117; PRO-1137 AND PHE-1589, AND VARIANT  
 RP ALA-1148.  
 RX MEDLINE=94108431; PubMed=8281141;  
 RA Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,  
 RA Berg M.A., Miller D.C., Francke U.;  
 RT "Mutation screening of complete fibrillin-1 coding sequence: report  
 RT of five new mutations, including two in 8-cysteine domains.",  
 RN Hum. Mol. Genet. 2:1813-1821(1993).  
 RN [18]  
 RP VARIANTS MFS GLY-217 AND ARG-2627.  
 RX MEDLINE=95067970; PubMed=7977366;  
 RA Karttunen L., Raghunath M., Lonnqvist L., Peltonen L.;

"A compound-heterozygous Marfan patient: two defective fibrillin  
 RT alleles result in a lethal phenotype.",  
 RT Am. J. Hum. Genet. 55:1083-1091(1994).  
 RN [19]  
 RP VARIANT EL LYS-2447.  
 RX MEDLINE=94245349; PubMed=8188302;  
 RA Lonnqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,  
 RA Peltonen L.;  
 RT "A novel mutation of the fibrillin gene causing ectopia lentis.",  
 RN Genomics 15:573-576(1994).  
 RN [20]  
 RP VARIANT MFS CYS-627.  
 RX MEDLINE=94272487; PubMed=8004112;  
 RA Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;  
 RT "Two novel mutations and a neutral polymorphism in EGF-like domains  
 RT of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan  
 RT syndrome patients.",  
 RN Hum. Mol. Genet. 3:373-375(1994).  
 RN [21]  
 RP VARIANT MFS CYS-122.  
 RX MEDLINE=94314977; PubMed=8040326;  
 RA Stahl-Hallgren C., Ukkonen T., Kainulainen K., Kristofersson U.,  
 RA Saxne T., Tornqvist K., Peltonen L.;  
 RT "An extra cysteine in one of the non-calcium-binding epidermal growth  
 RT factor-like motifs of the FBN1 polypeptide is connected to a novel  
 RT variant of Marfan syndrome.",  
 RN J. Clin. Invest. 94:709-713(1994).  
 RN [22]  
 RP VARIANT MFS TYR-1223.  
 RX MEDLINE=94351682; PubMed=8071963;  
 RA Hewett D.R., Lynch J.R., Child A., Sykes B.C.;  
 RT "A new missense mutation of fibrillin in a patient with Marfan  
 RT syndrome.",  
 RN J. Med. Genet. 31:338-339(1994).  
 RN [23]  
 RP VARIANT MFS HIS-1170.  
 RX MEDLINE=95174777; PubMed=7870075;  
 RA Hayward C., Porteous M.E.M., Brock D.J.H.;  
 RT "A novel mutation in the fibrillin gene (FBN1) in familial  
 RT arachnodactyly.",  
 RN Mol. Cell. Probes 8:325-327(1994).  
 RN [24]  
 RP VARIANTS MFS GLY-217; ASN-1023; ARG-1074; TYR-1242; ARG-1513;  
 RP GLU-2127; TRP-2151; LYS-2447 AND ARG-2511.  
 RX MEDLINE=94184368; PubMed=8136837;  
 RA Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;  
 RT "Mutations in the fibrillin gene responsible for dominant ectopia  
 RT lentis and neonatal Marfan syndrome.",  
 RN Nat. Genet. 6:64-69(1994).  
 RN [25]

Query Match 12.8%; Score 363; DB 1; Length 2871;  
 Best Local Similarity 30.8%; Pred. No. 1.9e-16;  
 Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;  
 QY 226 DCSVENGGCEHACNAIPGAPROCCPAGAAALQDGRSCTASATQSCNDLCEHFVCPNP---282  
 Db 1200 ECSIMNGGCTCTNTSEGSYEGSCQPGFALMPDQSRCT-----DIDE--CEDNPNC 1249  
 QY 283 -----DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPC-PORCVNTQGFECHCYP 335  
 Db 1250 DGGQCTNTPGEYRCLCYGFMASEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDM 1309  
 QY 336 NYDLVDGE--CPEVPDPC--FRANCEYQCPQLNQT-SYLCVCAEGFA-----PIP 380  
 Db 1310 GYSKGKGTCTD-INCEIGHNCGKHAIVCTNTAGSFACSCSPGWDGKICTDLDEGS 1368  
 QY 381 HEPHRCQMFQNOTACPADCPNQTAS--CECEGYILDDGFCTCTDIDEC-EN-----429  
 Db 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY-TGGDGFCTDLDCESENLNCGNGQ 1419  
 QY 430 -----GGF-----CS-----GVCHNLPGFTECIGPDSALV 455





modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U69263; AAC51260.2; -  
 DR EMBL; AL137638; CAB70853.1; ALT\_INIT.  
 DR HSSP; P01132; IEGF.  
 DR Genew; HGNC:6908; MATN2.  
 DR MIM; 602108; -  
 DR GO; GO:0005578; C:extracellular matrix; NAS.  
 DR InterPro; IPR000152; Asx hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF00008; EGF; 10.  
 DR Pfam; PF00092; vwa; 2.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00327; VWA\_2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 9.  
 DR PROSITE; PS01186; EGF\_2; 9.  
 DR PROSITE; PS50026; EGF\_3; 4.  
 DR PROSITE; PS50234; VWFA\_2.  
 DR EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil;  
 KW Alternative splicing.  
 FT SIGNAL 1 23  
 FT CHAIN 24 956  
 FT DOMAIN 57 232  
 FT DOMAIN 238 278  
 FT DOMAIN 279 319  
 FT DOMAIN 320 360  
 FT DOMAIN 361 401  
 FT DOMAIN 402 442  
 FT DOMAIN 443 483  
 FT DOMAIN 484 524  
 FT DOMAIN 525 565  
 FT DOMAIN 566 606  
 FT DOMAIN 607 647  
 FT DOMAIN 655 830  
 FT DOMAIN 917 955  
 FT DISULFID 242 253  
 FT DISULFID 249 262  
 FT DISULFID 264 277  
 FT DISULFID 283 294  
 FT DISULFID 290 303  
 FT DISULFID 305 318  
 FT DISULFID 324 335  
 FT DISULFID 331 344  
 FT DISULFID 346 359  
 FT DISULFID 365 376  
 FT DISULFID 372 385  
 FT DISULFID 387 400  
 FT DISULFID 406 417  
 FT DISULFID 413 426  
 FT DISULFID 428 441  
 FT DISULFID 447 458  
 FT DISULFID 454 467  
 FT DISULFID 469 482  
 FT DISULFID 488 499  
 FT DISULFID 495 508  
 FT DISULFID 510 523  
 FT DISULFID 529 540  
 FT DISULFID 536 549  
 FT DISULFID 551 564  
 FT DISULFID 570 581  
 FT DISULFID 577 590  
 FT DISULFID 592 605  
 FT DISULFID 611 622  
 FT DISULFID 618 631  
 FT DISULFID 633 646  
 FT CARBOHYD 221 221  
 FT CARBOHYD 890 890  
 FT VARSPLIC 861 879

FT CONFLICT 356 356 /FTid=VSP\_001399.  
 FT CONFLICT 594 594 K -> E (IN REF. 3).  
 FT CONFLICT 644 644 V -> E (IN REF. 3).  
 FT CONFLICT 755 755 R -> G (IN REF. 2).  
 FT CONFLICT 935 935 F -> L (IN REF. 3).  
 FT CONFLICT 935 935 L -> F (IN REF. 2).  
 SQ SEQUENCE 956 AA; 106840 MW; 826B7F347178FC80 CRC64;  
 Query Match 12.7%; Score 358; DB 1; Length 956;  
 Best Local Similarity 31.6%; Pred. No. 1.4e-16;  
 Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;  
 QY 227 CSVGNGGCEHACNAIPGAPRCQCPAGALQADGRSCTA-SATQSCNDLCEHFCVNPDPQ 285  
 DB 283 CAMEDHNCEQLCVNVPGSFVCCYSGVALAEDGKRCVAVDYCASNHGCEHCV-NAD-- 339  
 QY 286 GSYSCMCETGYRLAADQHRCEVDVDCILLEPSPQRCQVNTQGGFECHCYPNYDL-VDGEC 344  
 DB 340 GSYLCQCEHGFALNPDKTKTKIDYCASSNHGCOHECVNTDDSYSCHLKGFTLNPDKKT 399  
 QY 345 VEPVDPGF--RANCEYOCQPLNQTSYLCVCAEGFAPIPH-----EPHRCOMFC 390  
 DB 400 CRRINYCALNKPGEHCVCVNMEE-SYVCRCHRGYTLDPNGKTCRVVDHCAQDHGCCQLC 458  
 QY 391 NQTACPADCDPNTQAS--CECPGYILDDGF-ICTDIDEC---ENGFCGSGVCHNLPETF 444  
 DB 459 -----LNTEDSFVCCSGFLINEDLKTCSRVDYCLLSDHG--CEYSCVNMDSRF 506  
 QY 445 ECICGPDLSALVRHIGTDCDSKVDG---GDSG 473  
 DB 507 ACQC-PGHVLRSDGKTC--AKLDSALGDHG 535  
 RESULT 15  
 FBL2 MOUSE  
 ID FBL2 MOUSE STANDARD; PRT; 1221 AA.  
 AC P37889; O9WU12;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 GN Fibulin-2 precursor.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=94064787; PubMed=8245130;  
 RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;  
 RT "Structure and expression of fibulin-2, a novel extracellular matrix  
 RT protein with multiple EGF-like repeats and consensus motifs for  
 RT calcium binding.";  
 RL J. Cell Biol. 123:1269-1277(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=99337886; PubMed=10406956;  
 RA Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;  
 RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter  
 RT characterization.";  
 RL Eur. J. Biochem. 263:471-477(1999).  
 RN [3]  
 RP DEVELOPMENTAL STAGE.  
 RX MEDLINE=97003230; PubMed=8850569;  
 RA Zhang H.-Y., Timpl R., Sasaki T., Chu M.-L., Ekblom P.;  
 RT "Fibulin-1 and fibulin-2 expression during organogenesis in the  
 RT developing mouse embryo.";  
 RL Dev. Dyn. 205:348-364(1996).  
 RN [4]  
 RP BINDING TO LAMA2.  
 RX MEDLINE=99146904; PubMed=10022829;  
 RA Talts J.F., Andac Z., Goehring W., Brancaccio A., Timpl R.;

"Binding of the G domains of laminin alpha1 and alpha2 chains and perlecan to heparin, sulfatides, alpha-dystroglycan and several extracellular matrix proteins.";  
 RL EMBO J. 18:863-870(1999).

RT [5]  
 RP DOWN-REGULATION BY GLUCOCORTICOID.  
 RX MEDLINE=21600963; PubMed=11737251;  
 RA Gu Y.-C., Tals J.P., Gullberg D., Timpl R., Ekblom M.;  
 RT "Glucocorticoids down-regulate the extracellular matrix proteins  
 RL fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";  
 CC Eur. J. Haematol. 67:176-184(2001).  
 CC -!- FUNCTION: Its binding to fibronectin and some other ligands is  
 CC calcium dependent.  
 CC -!- SUBUNIT: Homotrimer; disulfide-linked. Interacts with LAMA2.  
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Comment-Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=P37889-1; Sequence=Displayed;  
 CC Name=2; Synonym=EGF3-less;  
 CC IsoId=P37889-2; Sequence=VSP\_001391;  
 CC -!- TISSUE SPECIFICITY: Component of both basement membranes and other  
 CC connective tissues.  
 CC -!- DEVELOPMENTAL STAGE: The differential expression of the fibulin  
 CC family contributes to the formation of molecularly distinct  
 CC extracellular matrices already during early developmental stages  
 CC of a large number of tissues.  
 CC -!- INDUCTION: Glucocorticoids suppressed mRNA expression and protein  
 CC synthesis.  
 CC -!- SIMILARITY: Belongs to the fibulin family.  
 CC -!- SIMILARITY: Contains 3 anaphylatoxin-like domains.  
 CC -!- SIMILARITY: Contains 11 EGF-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X75285; CAA53040.1; --  
 DR EMBL; AF135253; AAD34456.1; --  
 DR EMBL; AF135239; AAD34456.1; JOINED.  
 DR EMBL; AF135240; AAD34456.1; JOINED.  
 DR EMBL; AF135241; AAD34456.1; JOINED.  
 DR EMBL; AF135242; AAD34456.1; JOINED.  
 DR EMBL; AF135243; AAD34456.1; JOINED.  
 DR EMBL; AF135244; AAD34456.1; JOINED.  
 DR EMBL; AF135245; AAD34456.1; JOINED.  
 DR EMBL; AF135246; AAD34456.1; JOINED.  
 DR EMBL; AF135247; AAD34456.1; JOINED.  
 DR EMBL; AF135248; AAD34456.1; JOINED.  
 DR EMBL; AF135249; AAD34456.1; JOINED.  
 DR EMBL; AF135250; AAD34456.1; JOINED.  
 DR EMBL; AF135251; AAD34456.1; JOINED.  
 DR EMBL; AF135252; AAD34456.1; JOINED.  
 DR PIR; A49457; A49457.  
 DR HSSP; P00736; LAPO.  
 DR MGD; MGI:95488; Fbln2.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR000152; Asx\_hydroxyl\_s.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_like.  
 DR Pfam; PF01821; ANATO; 2.  
 DR Pfam; PF00008; EGF; 5.  
 DR SMART; SM00104; ANATO; 3.  
 DR SMART; SM00179; EGF\_Ca; 9.  
 DR PROSITE; PS01177; ANAPHYLATOXIN 1; 3.  
 DR PROSITE; PS01178; ANAPHYLATOXIN 2; 3.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
 DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF 2; 5.  
 DR PROSITE; PS00026; EGF 3; 5.  
 DR PROSITE; PS01187; EGF CA; 10.  
 KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
 KW Calcium-binding; Alternative splicing; Repeat.  
 FT SIGNAL 1 26  
 FT CHAIN 27 1221 FIBULIN-2.  
 FT DOMAIN 27 434 N.  
 FT DOMAIN 27 176 SUBDOMAIN NA (CYS-RICH).  
 FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).  
 FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 1.  
 FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 2.  
 FT DOMAIN 511 543 ANAPHYLATOXIN-LIKE 3.  
 FT DOMAIN 594 635 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 669 708 EGF-LIKE 2.  
 FT DOMAIN 709 755 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 756 800 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 801 846 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 847 894 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 895 937 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 938 979 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 980 1018 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1019 1061 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1062 1106 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1111 1221 DOMAIN III.  
 FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 435 462 BY SIMILARITY.  
 FT DISULFID 436 469 BY SIMILARITY.  
 FT DISULFID 449 470 BY SIMILARITY.  
 FT DISULFID 479 508 BY SIMILARITY.  
 FT DISULFID 492 509 BY SIMILARITY.  
 FT DISULFID 511 535 BY SIMILARITY.  
 FT DISULFID 512 542 BY SIMILARITY.  
 FT DISULFID 525 543 BY SIMILARITY.  
 FT DISULFID 598 610 BY SIMILARITY.  
 FT DISULFID 606 619 BY SIMILARITY.  
 FT DISULFID 621 634 BY SIMILARITY.  
 FT DISULFID 673 683 BY SIMILARITY.  
 FT DISULFID 679 692 BY SIMILARITY.  
 FT DISULFID 694 707 BY SIMILARITY.  
 FT DISULFID 713 726 BY SIMILARITY.  
 FT DISULFID 720 735 BY SIMILARITY.  
 FT DISULFID 742 754 BY SIMILARITY.  
 FT DISULFID 805 818 BY SIMILARITY.  
 FT DISULFID 812 827 BY SIMILARITY.  
 FT DISULFID 833 845 BY SIMILARITY.  
 FT DISULFID 899 912 BY SIMILARITY.  
 FT DISULFID 906 921 BY SIMILARITY.  
 FT DISULFID 923 936 BY SIMILARITY.  
 FT DISULFID 942 954 BY SIMILARITY.  
 FT DISULFID 950 963 BY SIMILARITY.  
 FT DISULFID 965 978 BY SIMILARITY.  
 FT DISULFID 984 993 BY SIMILARITY.  
 FT DISULFID 989 1002 BY SIMILARITY.  
 FT DISULFID 1004 1017 BY SIMILARITY.  
 FT DISULFID 1023 1035 BY SIMILARITY.  
 FT DISULFID 1031 1044 BY SIMILARITY.  
 FT DISULFID 1046 1060 BY SIMILARITY.  
 FT DISULFID 1065 1079 BY SIMILARITY.  
 FT DISULFID 1073 1088 BY SIMILARITY.  
 FT DISULFID 1093 1105 BY SIMILARITY.  
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 709 755 Missing (in isoform 2).  
 FT FTId=VSP\_001391.  
 FT HSGRYAGHTVHLSCRCAC -> TVAVSICWYRPPLILP  
 FT GF (IN REF. 2).  
 FT S -> L (IN REF. 2).  
 FT Q -> QQ (IN REF. 2).  
 FT Q -> E (IN REF. 2).  
 FT CONFLICT 1102 1102  
 FT CONFLICT 1221 AA; 131818 MW; 87DBZAL0A8FDC45F CRC64;  
 FT SEQUENCE



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:43:13 ; Search time 37.5 seconds  
(without alignments)  
4190.082 Million cell updates/sec

Title: US-09-509-994-1\_COPY\_19\_516  
Perfect score: 2830  
Sequence: 1 APAEPQGGSCQVHDCFAL.....PSTPGSLTPPAVLVHSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvrius:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 2830   | 100.0       | 625    | 4     | Q81V29      |
| 2          | 2654   | 93.8        | 468    | 4     | Q9UC32      |
| 3          | 1983.5 | 70.1        | 580    | 6     | Q8H248      |
| 4          | 1786   | 63.1        | 577    | 11    | O35370      |
| 5          | 1776.5 | 62.8        | 491    | 11    | Q8BJB5      |
| 6          | 1396.5 | 49.3        | 461    | 11    | P97883      |
| 7          | 574    | 20.3        | 652    | 4     | Q81XK1      |
| 8          | 574    | 20.3        | 757    | 4     | Q9HCU0      |
| 9          | 566.5  | 20.0        | 765    | 11    | O91ZV1      |
| 10         | 566.5  | 20.0        | 765    | 11    | Q91V98      |
| 11         | 433    | 15.3        | 1654   | 5     | Q9TVQ2      |
| 12         | 414    | 14.6        | 1574   | 11    | O88281      |
| 13         | 385    | 13.6        | 1409   | 5     | Q9V889      |
| 14         | 375    | 13.3        | 708    | 13    | P87363      |
| 15         | 374    | 13.2        | 2906   | 11    | Q9WUH9      |
| 16         | 369    | 13.0        | 1231   | 4     | Q81U11      |

|    |       |      |      |    |        |                     |
|----|-------|------|------|----|--------|---------------------|
| 17 | 369   | 13.0 | 1246 | 4  | O75095 | O75095 homo sapien  |
| 18 | 367.5 | 13.0 | 1184 | 4  | Q86V58 | Q86V58 homo sapien  |
| 19 | 366   | 12.9 | 3857 | 11 | O88840 | O88840 mus musculus |
| 20 | 365   | 12.9 | 2809 | 4  | Q96JP8 | Q96JP8 homo sapien  |
| 21 | 365   | 12.9 | 2809 | 4  | Q86SJ5 | Q86SJ5 homo sapien  |
| 22 | 360   | 12.7 | 528  | 11 | Q9CXD8 | Q9CXD8 mus musculus |
| 23 | 360   | 12.7 | 2872 | 11 | Q9WUH8 | Q9WUH8 rattus norv  |
| 24 | 359   | 12.7 | 715  | 4  | Q7Z5X1 | Q7Z5X1 homo sapien  |
| 25 | 359   | 12.7 | 937  | 4  | Q96FT5 | Q96FT5 homo sapien  |
| 26 | 359   | 12.7 | 1016 | 4  | Q8NDE6 | Q8NDE6 homo sapien  |
| 27 | 358.5 | 12.7 | 893  | 6  | Q8MJK0 | Q8MJK0 cercopithec  |
| 28 | 358.5 | 12.7 | 1174 | 11 | Q99K58 | Q99K58 mus musculus |
| 29 | 357   | 12.6 | 590  | 11 | Q8C088 | Q8C088 mus musculus |
| 30 | 356   | 12.6 | 1231 | 4  | Q81U10 | Q81U10 homo sapien  |
| 31 | 353.5 | 12.5 | 746  | 4  | Q96HB9 | Q96HB9 homo sapien  |
| 32 | 353.5 | 12.5 | 1256 | 4  | Q9NS15 | Q9NS15 homo sapien  |
| 33 | 353.5 | 12.5 | 1382 | 4  | Q9H7K2 | Q9H7K2 homo sapien  |
| 34 | 352.5 | 12.5 | 741  | 4  | Q96K89 | Q96K89 homo sapien  |
| 35 | 351.5 | 12.4 | 1511 | 4  | O75412 | O75412 homo sapien  |
| 36 | 351.5 | 12.4 | 1587 | 4  | O00508 | O00508 homo sapien  |
| 37 | 351   | 12.4 | 1666 | 11 | Q8K4G1 | Q8K4G1 mus musculus |
| 38 | 350   | 12.4 | 1600 | 11 | Q8K4G0 | Q8K4G0 mus musculus |
| 39 | 346.5 | 12.2 | 729  | 11 | Q8BNH3 | Q8BNH3 mus musculus |
| 40 | 346.5 | 12.2 | 787  | 11 | Q8K061 | Q8K061 mus musculus |
| 41 | 344.5 | 12.2 | 589  | 13 | Q7ZX63 | Q7ZX63 xenopus lae  |
| 42 | 343   | 12.1 | 5636 | 4  | Q96RW7 | Q96RW7 homo sapien  |
| 43 | 341   | 12.0 | 576  | 4  | Q9Y3V7 | Q9Y3V7 homo sapien  |
| 44 | 338   | 11.9 | 1062 | 11 | Q60789 | Q60789 mus musculus |
| 45 | 337.5 | 11.9 | 1289 | 5  | Q86AC3 | Q86AC3 dictyosteli  |

## ALIGNMENTS

## RESULT 1

Q81V29  
ID Q81V29 PRELIMINARY; PRT; 625 AA.  
AC Q81V29  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombomodulin (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035602; AAH35602.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005229; F:sugar binding; IEA.  
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.  
DR GO; GO:0007596; P:blood coagulation; IEA.  
DR InterPro; IPR000152; ASX hydroxyl\_s.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR001304; Lectin\_C.  
DR InterPro; IPR001491; Thrombomoduln.  
DR Pfam; PF00008; EGF; 3.  
DR Pfam; PF00059; lectin\_c; 1.  
DR PRINTS; PR00907; THROMBOMODULN.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 6.  
DR SMART; SM00179; EGF\_Ca; 4.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS00041; C TYPE LECTIN\_2; 1.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS01187; EGF\_Ca; 2.



```

FT NON TER      1      1
SQ SEQUENCE 625 AA; 65876 MW; 3E5E161183A854B9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 2830; DB 4; Length 625;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
DB 69 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 128
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLPLRGFQWVTGDNNTSYSRWALDLNGLAPLGGPLCV 120
DB 129 GGVGRRRLWIGLQPLPGCGDPKRLPLRGFQWVTGDNNTSYSRWALDLNGLAPLGGPLCV 188
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTGPPA 180
DB 189 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTGPPA 248
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
DB 249 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 308
QY 241 IPGAPRCQCPAGAAALOADGRSCTASATQSCNDLCEHFQVNPDPGSGYSCMCTGYRLAA 300
DB 309 IPGAPRCQCPAGAAALOADGRSCTASATQSCNDLCEHFQVNPDPGSGYSCMCTGYRLAA 368
QY 301 DQHRCEVDVDCILPEPSPQPCQVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
DB 369 DQHRCEVDVDCILPEPSPQPCQVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 428
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYLLDDGFI 420
DB 429 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYLLDDGFI 488
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGEPPPS 480
DB 489 CTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVDGSGSGEPPPS 548
QY 481 PTPGSTLTTPPAVGLVHSG 498
DB 549 PTPGSTLTTPPAVGLVHSG 566

RESULT 2
Q9UC32 PRELIMINARY; PRT; 468 AA.
AC Q9UC32;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombomodulin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=93293792; PubMed=8390446;
RX Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;
RA "Urinary thrombomodulin, its isolation and characterization.";
RL J. Biochem. 113:433-440(1993).
DR HSP: P07204; IZAK.
DR GO: 0016020; C:membrane; IEA.
DR GO: 0005509; F:calcium ion binding; IEA.
DR GO: 0005529; F:sugar binding; IEA.
DR GO: 0004888; F:transmembrane receptor activity; IEA.
DR GO: 0007596; P:blood coagulation; IEA.
DR InterPro: IPR00152; Asx_hydroxyl_S.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR001304; Lectin_C.

```

```

DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_3.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 2.
KW EGF-like domain.
SQ SEQUENCE 468 AA; 49444 MW; 4BFE8E98EBFB6A40 CRC64;

Query Match
Best Local Similarity 99.8%; Score 2654; DB 4; Length 468;
Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
DB 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLPLRGFQWVTGDNNTSYSRWALDLNGLAPLGGPLCV 120
DB 61 GGVGRRRLWIGLQPLPGCGDPKRLPLRGFQWVTGDNNTSYSRWALDLNGLAPLGGPLCV 120
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTGPPA 180
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTGPPA 180
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
QY 241 IPGAPRCQCPAGAAALOADGRSCTASATQSCNDLCEHFQVNPDPGSGYSCMCTGYRLAA 300
DB 241 IPGAPRCQCPAGAAALOADGRSCTASATQSCNDLCEHFQVNPDPGSGYSCMCTGYRLAA 300
QY 301 DQHRCEVDVDCILPEPSPQPCQVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
DB 301 DQHRCEVDVDCILPEPSPQPCQVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYLLDDGFI 420
DB 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYLLDDGFI 420
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVD 468
DB 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDCSGKVD 468

RESULT 3
Q8HZ48 PRELIMINARY; PRT; 580 AA.
AC Q8HZ48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombomodulin precursor.
GN THBD.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX Deming C.B., Kim A.Y., Rade J.J.;
RA "Cloning of rabbit thrombomodulin cDNA sequence.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AV138902; AAM15931.2; -.
DR FIR: A33308; A33308.
DR GO: 0016020; C:membrane; IEA.
DR GO: 0005509; F:calcium ion binding; IEA.
DR GO: 0005529; F:sugar binding; IEA.

```

[illegible]

## RESULT 4

| ID      | PRELIMINARY;  | PET;                                    | 577 AA. |
|---------|---|---|---------|
| Q35370; |   |   |         |
| AC      |   |   |         |
| DI      | Q35370; 1998  | (T=EMBLrel. 05, Created)                |         |
| DT      | 01-JAN-1998   | (T=EMBLrel. 05, Last sequence update)   |         |
| DT      | 01-OCT-2003   | (T=EMBLrel. 25, Last annotation update) |         |
| DE      | Thrombosmodulin.  |   |         |
| OS      | Rattus norvegicus (Rat).  |   |         |
| OC      | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |         |
| OC      | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.      |   |         |

NCBI\_TaxID=10116;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;  
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE OF 1-136 FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF022743; AAB80760.1; -;  
DR EMBL; AF022742; AAB80923.1; -;  
DR HSSP; P07204; 1FGD.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005529; F:sugar binding; IEA.  
DR GO; GO:0004488; F:transmembrane receptor activity; IEA.  
DR GO; GO:0007536; P:blood coagulation; IEA.  
DR InterPro; IPRO00152; Asx hydroxyl\_S.  
DR InterPro; IPRO01881; EGF\_Ca.  
DR InterPro; IPRO06209; EGF\_like.  
DR InterPro; IPRO01304; Lectin\_C.  
DR InterPro; IPRO01491; Thrombomodulin.  
DR Pfam; PF00008; EGF; 3.  
DR Pfam; PF00059; lectin c; 1.  
DR PRINTS; PR00307; THROMBOMODULIN.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00179; EGF\_CA; 2.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS00041; C TYPE LECTIN\_2; 1.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01187; EGF\_CA; 2.  
DR EGF-like domain.  
SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;

| Query Match           | 63.11% | Score 1786;         | DB 11;          | Length 577; |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
|-----------------------|--------|---------------------|-----------------|-------------|--------|------|----|------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|-----|-----|
| Best Local Similarity | 63.74% | Pred. No. 9.7e-140; |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Matches               | 320;   | Conservative 48;    | Mismatches 128; | Indels 6;   | Gaps 5 |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 1      | APASEPQGGSCV        | HD              | CFAL        | VP     | GATF | LN | ASQI | CD | GL | RH | LM | TV | RS | SS | VA | AD | VI | S  | LL | NG | 60 |    |     |     |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 19     | ALAKLPKGS           | CV              | GN          | EC     | FAL  | QD | PV   | TF | LD | AS | QA | CQ | RL | CH | LM | TV | RS | SS | VA | AD | 77 |    |     |     |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 61     | GGVCRRL             | W               | I           | GL     | QL   | P  | PG   | CG | D  | K  | R  | GL | P  | RG | FW | MT | GD | NN | TS | YS | RW | LD | 120 |     |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 78     | SSMDSRP             | -               | W           | I      | GL   | QL | P    | Q  | CG | D  | V  | H  | LG | P  | RG | FW | MT | GD | NN | TS | YS | RW | LD  | 136 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 121    | AVSAAEATV           | PS              | EP          | T      | W    | E  | E    | Q  | Q  | C  | K  | V  | K  | AD | G  | F  | L  | C  | E  | H  | F  | F  | AT  | 179 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 137    | TVSTAT              | EAA             | P           | G      | E    | A  | W    | E  | K  | P  | C  | E  | N  | E  | T  | K  | G  | F  | L  | C  | E  | F  | F   | AT  | 196 |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 180    | AARGAD              | FO              | AL          | P      | V    | G  | S    | A  | A  | V  | A  | P  | L  | G  | L  | M  | C  | T  | A  | P  | P  | G  | A   | 239 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 197    | GVSGAD              | F               | Q           | T      | L    | P  | I    | G  | S  | S  | A  | T  | V  | A  | P  | L  | G  | L  | V  | C  | R  | A  | L   | P   | 256 |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 240    | AI                  | PC              | A           | R      | C    | Q  | C    | P  | A  | G  | A  | L  | O  | A  | D  | G  | S  | C  | T  | A  | S  | A  | T   | 299 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 257    | RS                  | AN              | P           | R      | C    | V  | C    | P  | S  | G  | D  | L  | O  | A  | D  | G  | S  | C  | A  | K  | P  | V  | A   | 316 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 300    | AD                  | H               | R           | C      | E    | D  | V    | D  | C  | I  | L  | E  | P  | S  | P  | C  | P  | Q  | R  | C  | V  | N  | T   | 359 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Db                    | 317    | AD                  | H               | R           | C      | E    | D  | V    | D  | C  | K  | O  | G  | N  | P  | C  | P  | Q  | L  | C  | N  | T  | E  | G   | 376 |     |
|                       |        |                     |                 |             |        |      |    |      |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |     |     |
| Qy                    | 360    | CQ                  | P               | N           | Q      | T    | S  | L    | C  | V  | A  | E  |    |    |    |    |    |    |    |    |    |    |    |     |     |     |

```
Db 497 PSSNPTVSVSTVPSARPMSG 518
Q8BJB5
ID Q8BJB5 PRELIMINARY; PRT; 491 AA.
AC Q8BJB5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombomodulin precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK089479; BAC40898.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_3.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF_6.
DR SMART; SM00179; EGF_CA; 4.
DR PROSITE; PS00010; ASX HYDROXYL; 2.
DR PROSITE; PS00041; C-TYPE LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 2.
FT NON TER 491
SQ SEQUENCE 491 AA; 52904 MW; AE2007BEC68AD20D CRC64;

Query Match 62.8%; Score 1776.5; DB 11; Length 491;
Best Local Similarity 67.1%; Pred. No. 5e-139;
Matches 312; Conservative 43; Mismatches 105; Indels 5; Gaps 3;

Qy 1 APAEPQGGSCVHDCFALYPGPATFINASQICDGLRGLMTVRSSVAADVISLLNGD 60
Db 19 ALAKLQPTGSCVHECFALFQGPATFLDASQAQRLOGHLMTVRSSVAADVISLLSQS 78
Qy 61 G-GVGRRLWIGLQPLPGCGDPKELGRLGHQWTGDNNTSYSWARLDLNGAPLCGPLC 119
Db 79 SMDLGP--WIGLQPLGCGDDPVHLGLRGLRQWVTGDNHTSYSWARPNDQAPLCGPLC 135
Qy 120 VAVSAAEATVPSEPIWEEQCEVKADGLCFHFHFPATCRPLAVEP-GAAAAVSITYGTP 178
Db 136 VTVSTATAEARGEFAWEKPCETETQGLCFEYFTASCRPTVTNTRDPEAAHSSITYNP 195
Qy 179 FAARGADFQALPVGSSAAVAPLGLQMCATPPGAVQGHWAREAPGAWDCSVENGCCSHAC 238
Db 196 FGVSGADFQTLFVGSSAAVEPLGLVLCRAPPGTSEGHWAREATGANWCSVENGCCBYLC 255
Qy 239 NAIFGARPCQCPAGAAALQADGRSCTASATQSCNDLCEHFCHYCNVLDVGCSEVPDVP 298
Db 256 NRSTNEPRCLPRDMDLQADGRSCARPVVQSCNLCCEHFCHYCNVLDVGCSEVPDVP 315
Qy 299 AADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYCNVLDVGCSEVPDVPDPCFRANCEY 358
Db 316 AADQHRCEVDVDCCKQGNPCQRCVNTQGGFECCHYCNVLDVGCSEVPDVPDPCFRANCEY 375

Qy 359 QCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTQACPADCDNTQASCECPGYYILDG 418
Db 376 QCOQVFTDYRCICAPGFAPKPDPEPHKCFMFCNETSCPADCDNSPTVCPCPGFILDG 435
Qy 419 FICTIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCD 463
Db 436 SVCTIDECESQCECTSECFNFCGYSYECICGPDPTALAGQISKDCD 480

RESULT 6
P97883
ID P97883 PRELIMINARY; PRT; 461 AA.
AC P97883;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thrombomodulin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain capillary;
RA Wang L., Tran N.D., Schreiber S.S., Zickovic B.V.;
RT "Nucleotide sequence of rat thrombomodulin.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90121; AAB49723.1; -.
DR HSPB; P07204; 1AFGD.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007596; F:blood coagulation; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF_3.
DR PRINTS; PR00907; THROMBOMODULIN.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00010; ASX HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 2.
DR EGF-like domain.
FT NON TER 1
SQ SEQUENCE 461 AA; 49499 MW; 45971BCE8468B67 CRC64;

Query Match 49.3%; Score 1396.5; DB 11; Length 461;
Best Local Similarity 62.4%; Pred. No. 1.5e-107;
Matches 244; Conservative 37; Mismatches 105; Indels 5; Gaps 4;

Qy 113 PLGCLPCLVAVSAAEATVPSEPIWEEQCEVKADGLCFHFHFPATC-RPLAVEP-GAAAA 170
Db 12 PLGCLPCLVTVSTATAEARGEFAWEKPCENETKGLCFEYFAAFCSRPLRVNTRDPEGAH 71
Qy 171 VSTYTGTPFAARCAADQALPVGSSAAVAPLGLQMCATPPGAVQGHWAREAPGAWDCSV 230
Db 72 ISSTYTVPLGVSGADFTPLIGSSATVAPFGLVLCRALPGTSEGHWAREAPGAWDCSV 131
Qy 231 NGGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCHYCNVLDVGCSEV 290
Db 132 NGGCEYMCNRSANGPRCVCPSGGDLQADGRSCKAPVQGLCNELCOHFVNNSDVPGSYSC 191
Qy 291 MCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYCNVLDVGCSEVPDVP 350
Db 192 MCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHYCNVLDVGCSEVPDVP 251
Qy 351 CFRANCEYCOPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTQACPADCDNTQASCECP 410
Db 252 CFRSKEYCOQPNVSTHYNCICAEFGAPKLDDPRCEMFNCTSCPADCDNSPFCQCP 311
Qy 411 EGYLDDGFTCTDIDECENGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDCDCKV--D 468
```

```

Db 312 EGFIDEGSICTDIDEGSQECLTNECRNLPGSYECIGPDTALAGQISKDCDPIPVLED 371
QY 469 GDSGSGEPPPS-PTPGSTLTTPRAVLVHSG 498
Db 372 SEDGSGEHPSSNFTVVSSTVPPSPARPMHSG 402

RESULT 7
Q8IXK1
ID Q8IXK1 PRELIMINARY; PRT; 652 AA.
AC Q8IXK1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Clq receptor protein precursor.
GN CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steinberger P., Stoeckl J., Wille S., Szekeres A., Prager E.,
RA Staefler G., Kuenig S., Kohl P., Majdic O., Knapp W., Stockinger H.;
RT "Identification of CD93 as the Clq receptor protein (ClqR) by
RT retroviral expression cloning.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ295142; CAC82720.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IRGF.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR001187; Tissue_factor.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF01108; Tissue_fac; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00179; EGF_CA; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
DR Signal; Signal.
KW SIGNAL.
FT CHAIN 1 21 POTENTIAL.
FT CHAIN 22 652 C1Q RECEPTOR PROTEIN.
SQ SEQUENCE 652 AA; 68521 MW; 97549BA62CAF225E CRC64;

Query Match
Best Local Similarity 31.4%; Pred. No. 3,6e-39;
Matches 174; Conservative 52; Mismatches 202; Indels 126; Gaps 31;

QY 6 QPGSGQ-----CVHEPCALYPGDPATFINASQICDGLRGHLMTVRSSVAAD-----VI 53
Db 17 QPGAGTGADTEAVVCVGTACTYTAHSGKLSAAEQNHQNGNLATVKSKEAQHVQRVL 76
QY 54 SLILNGGVGRR--RLWIGLQLPPG-CGDPKRLGLRGLFQWVTGDNNTYSRWARLDLN 110
Db 77 AQLLRRAALTARMSKFWIGLQREKGLDPSL--PLKGFSGWGGEDTPTYSNNHKELRN 134
QY 111 GAPLCGFLCVA--VSAEAATVPS-PWEEQC-----EVKADGFLCEFFHFPATCRPLA 161
Db 135 SC--ISKRCVSLLDLSQPLLPSRLPKWSEGPCSGSPGSNIEGFFVCKFSKGMCHPLA 192
QY 162 V-EPGAAAIVSYTGTPTFAARGADFOALPVGSSAAVAPLGLQMC-TAPPGAVQGHW-- 217
Db 193 LGGFG-----QVTTYTTPFTQTSSSLEAVPFAAANVA-----CGEGDKDETQSHYFL 239

```

```

QY 218 -AREAPGAWD-----CSVENGCGEHAC-NAIPGAPRCOCAGAAALQADGRS 261
Db 240 CKEKAPDFVDMGSSGFLCVSPKYGCNFFNGGCHQDCFEFGDGSFLCGCRFRLDLVLT 299
QY 262 CTASATQSCNDLCE--HFCVNPDPQGSYSCMCETGYRLAADQHRCEVDVDCILLEPSPCP 319
Db 300 C-ASRNPCSSPCRGATCALGP-HGKNYTCRFGYQLDSSQLDQDVDEEC--QDSPCA 355
QY 320 QRCVNTQGGFECHYCNPNYDLVDGECVEPDPFCFRANCEYQCQPLNQTSYLCVCAEGFAP 379
Db 356 QECVNTPGGPRCEWGVY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395
QY 380 PHEPHRCQMECNQTACPADCPNPTQAS--CECPEGYL--DDGFICTDIDEC--ENGGFC 433
Db 396 -----CAQGC-TNTDGSFHCSCEGVVLAGEDGTQCQDVDECVGPGGFLC 439
QY 434 SGVCHNLPGTFECICGPDSDSALVRHIGTDCDSGKV-----DGGDSGSGE-----P 477
Db 440 DSLCFNTQGSFHCGLPGWVLAEN-GVSCVTMGVSLGPPSGPPDEEDKGEKSTVPRAA 498
QY 478 PPSPTPGSTLTTPPA 491
Db 499 TASPTRGPGEGTPKA 512

RESULT 8
Q9HCUC0
ID Q9HCUC0 PRELIMINARY; PRT; 757 AA.
AC Q9HCUC0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tumor endothelial marker 1 precursor (Endosialin protein).
GN TM1 OR ENDOSIALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium.";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT "Molecular cloning and characterization of Endosialin, a C-type
RT lectin-like cell surface receptor of tumor Endothelium.";
RL J. Biol. Chem. 276:7408-7414(2001).
DR EMBL; AF279142; ABG00867.1; -.
DR EMBL; AJ295846; CAC34381.1; -.
DR HSP; P07204; IZAQ.
DR Genew; HGNC:18219; CD164L1.
DR GO; GO:0005578; C:extracellular matrix; TAS.
DR InterPro; IPR000152; ASX_HYDROXYL_S.
DR InterPro; IPR001891; EGF_CA.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR EGF-like domain; Signal.
KW

```

[illegible]

RT "Molecular Characterization of the Mouse Tem1/endosomal Gene  
 RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in  
 RT Vivo.";

RL J. Biol. Chem. 276:38795-38807(2001).

RN [3]

RN SEQUENCE FROM N.A.

RC TISSUE=Olfactory epithelium;

RX MEDLINE=2238257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore S.I., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Scherch J.E.,  
 RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]

RN SEQUENCE FROM N.A.

RC TISSUE=Olfactory epithelium;

RX Strausberg R.J.

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF378758; AAL11995.1; -

DR EMBL; AF388573; AAK84665.1; -

DR EMBL; BC046318; AAK46318.1; -

DR MGD; MGI:1917695; Tem1.

DR GO; GO:0005737; Cytoplasm; IDA.

DR InterPro; IPR001152; ASX\_hydroxyl\_S.

DR InterPro; IPR001181; EGF\_CA.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR001304; Lectin\_C.

DR InterPro; IPR000436; Sushi\_SCR\_CCP.

DR Pfam; PF00008; EGF\_2.

DR Pfam; PF00059; Lectin\_c; 1.

DR Pfam; PF00084; sushi\_1.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00179; EGF\_CA; 1.

DR PROSITE; PS00010; ASX\_HYDROXYL; 1.

DR PROSITE; PS00041; C\_TYPE\_LECTIN\_2; 1.

DR PROSITE; PS01186; EGF\_2; 1.

DR PROSITE; PS01187; EGF\_CA; 1.

KW EGF-like domain; Signal.

FT SIGNAL

FT CHAIN

FT POTENTIAL.

FT 17 TUMOR ENDOTHELIAL MARKER 1.

FT 18 765

FT 19

FT 20

FT 21

FT 22

FT 23

FT 24

FT 25

FT 26

FT 27

FT 28

FT 29

FT 30

FT 31

FT 32

FT 33

FT 34

FT 35

FT 36

FT 37

FT 38

FT 39

FT 40

FT 41

FT 42

FT 43

FT 44

FT 45

FT 46

FT 47

FT 48

FT 49

FT 50

FT 51

FT 52

FT 53

FT 54

FT 55

FT 56

FT 57

FT 58

FT 59

FT 60

FT 61

FT 62

FT 63

FT 64

FT 65

FT 66

FT 67

FT 68

FT 69

FT 70

FT 71

FT 72

FT 73

FT 74

FT 75

FT 76

FT 77

FT 78

FT 79

FT 80

FT 81

FT 82

FT 83

FT 84

FT 85

FT 86

FT 87

FT 88

FT 89

FT 90

FT 91

FT 92

FT 93

FT 94

FT 95

FT 96

FT 97

FT 98

FT 99

FT 100

FT 101

FT 102

FT 103

FT 104

FT 105

FT 106

FT 107

FT 108

FT 109

FT 110

FT 111

FT 112

FT 113

FT 114

FT 115

FT 116

FT 117

FT 118

FT 119

FT 120

FT 121

FT 122

FT 123

FT 124

FT 125

FT 126

FT 127

FT 128

FT 129

FT 130

FT 131

FT 132

FT 133

FT 134

FT 135

FT 136

FT 137

FT 138

FT 139

FT 140

FT 141

FT 142

FT 143

FT 144

FT 145

FT 146

FT 147

FT 148

FT 149

FT 150

FT 151

FT 152

FT 153

FT 154

FT 155

FT 156

FT 157

FT 158

FT 159

FT 160

FT 161

FT 162

FT 163

FT 164

FT 165

FT 166

FT 167

FT 168

FT 169

FT 170

FT 171

FT 172

FT 173

FT 174

FT 175

FT 176

FT 177

FT 178

FT 179

FT 180

FT 181

FT 182

FT 183

FT 184

FT 185

FT 186

FT 187

FT 188

FT 189

FT 190

FT 191

FT 192

FT 193

FT 194

FT 195

FT 196

FT 197

FT 198

FT 199

FT 200

FT 201

FT 202

FT 203

FT 204

FT 205

FT 206

FT 207

FT 208

FT 209

FT 210

FT 211

FT 212

FT 213

FT 214

FT 215

FT 216

FT 217

FT 218

FT 219

FT 220

FT 221

FT 222

FT 223

FT 224

FT 225

FT 226

FT 227

FT 228

FT 229

FT 230

FT 231

FT 232

FT 233

FT 234

FT 235

FT 236

FT 237

FT 238

FT 239

FT 240

FT 241

FT 242

FT 243

FT 244

FT 245

FT 246

FT 247

FT 248

FT 249

FT 250

FT 251

FT 252

FT 253

FT 254

FT 255

FT 256

FT 257

FT 258

FT 259

FT 260

FT 261

FT 262

FT 263

FT 264

FT 265

FT 266

FT 267

FT 268

FT 269

FT 270

FT 271

FT 272

FT 273

FT 274

FT 275

FT 276

FT 277

FT 278

FT 279

FT 280

FT 281

FT 282

FT 283

FT 284

FT 285

FT 286

FT 287

FT 288

FT 289

FT 290

FT 291

FT 292

FT 293

FT 294

FT 295

FT 296

FT 297

FT 298

FT 299

FT 300

FT 301

FT 302

FT 303

FT 304

FT 305

FT 306

FT 307

FT 308

FT 309

FT 310

FT 311

FT 312

FT 313

FT 314

FT 315

FT 316

FT 317

FT 318

FT 319

FT 320

FT 321

FT 322

FT 323

FT 324

FT 325

FT 326

FT 327

FT 328

FT 329

FT 330

FT 331

FT 332

FT 333

FT 334

FT 335

FT 336

FT 337

FT 338

FT 339

FT 340

FT 341

FT 342

FT 343

FT 344

FT 345

FT 346

FT 347

FT 348

FT 349

FT 350

FT 351

FT 352

FT 353

FT 354

FT 355

FT 356

FT 357

FT 358

FT 359

FT 360

FT 361

FT 362

FT 363

FT 364

FT 365

FT 366

FT 367

FT 368

FT 369

FT 370

FT 371

FT 372

FT 373

FT 374

FT 375

FT 376

FT 377

FT 378

FT 379

FT 380

FT 381

FT 382

FT 383

FT 384

FT 385

FT 386

FT 387

FT 388

FT 389

FT 390

FT 391

FT 392

FT 393

FT 394

FT 395

FT 396

FT 397

FT 398

FT 399

FT 400

FT 401

FT 402

FT 403

FT 404

FT 405

FT 406

FT 407

FT 408

FT 409

FT 410

FT 411

FT 412

FT 413

FT 414

FT 415

FT 416

FT 417

FT 418

FT 419

FT 420

FT 421

FT 422

FT 423

FT 424

FT 425

FT 426

FT 427

FT 428

FT 429

FT 430

FT 431

FT 432

FT 433

FT 434

FT 435

FT 436

FT 437

FT 438

FT 439

FT 440

FT 441

FT 442

FT 443

FT 444

FT 445

FT 446

FT 447

FT 448

FT 449

FT 450

FT 451

FT 452

FT 453

FT 454

FT 455

FT 456

FT 457

FT 458

FT 459

FT 460

FT 461

FT 462

FT 463

FT 464

FT 465

FT 466

FT 467

FT 468

FT 469

FT 470

FT 471

FT 472

FT 473

FT 474

FT 475

FT 476

FT 477

FT 478

FT 479

FT 480

FT 481

FT 482

FT 483

FT 484

FT 485

FT 486

FT 487

FT 488

FT 489

FT 490

FT 491

FT 492

FT 493

FT 494

FT 495

FT 496

FT 497

FT 498

FT 499

FT 500

FT 501

FT 502

FT 503

FT 504

FT 505

FT 506

FT 507

FT 508

FT 509

FT 510

FT 511

FT 512

FT 513

FT 514

FT 515

FT 516

FT 517

FT 518

FT 519

FT 520

FT 521

FT 522

FT 523

FT 524

FT 525

FT 526

FT 527

FT 528

FT 529

FT 530

FT 531

FT 532

FT 533

FT 534

FT 535

FT 536

FT 537

FT 538

FT 539

FT 540

FT 541

FT 542

FT 543

FT 544

FT 545

FT 546

FT 547

FT 548

FT 549

FT 550

FT 551

FT 552

FT 553

FT 554

FT 555

FT 556

FT 557

FT 558

FT 559

FT 560

FT 561

FT 562

FT 563

FT 564

FT 565

FT 566

FT 567

FT 568

FT 569

FT 570

FT 571

FT 572

FT 573

FT 574

FT 575

FT 576

FT 577

FT 578

FT 579

FT 580

FT 581

FT 582

FT 583

FT 584

FT 585

FT 586

FT 587

FT 588

FT 589

FT 590

FT 591

FT 592

FT 593

FT 594

FT 595

FT 596

FT 597

FT 598

FT 599

FT 600

FT 601

FT 602

FT 603

FT 604

FT 605

FT 606

FT 607

FT 608

FT 609

FT 610

FT 611

FT 612

FT 613

FT 614

FT 615

FT 616

FT 617

FT 618

FT 619

FT 620

FT 621

FT 622

FT 623

FT 624

FT 625

FT 626

FT 627

FT 628

FT 629

FT 630

FT 631

FT 632

FT 633

FT 634

FT 635

FT 636

FT 637

FT 638

FT 639

FT 640

FT 641

FT 642

FT 643

FT 644

FT 645

FT 646

FT 647

FT 648

FT 649

FT 650

FT 651

FT 652

FT 653

FT 654

FT 655

FT 656

FT 657

FT 658

FT 659

FT 660

FT 661

FT 662

FT 663

FT 664

FT 665

FT 666

FT 667

FT 668

FT 669

FT 670

FT 671

FT 672

FT 673

FT 674

FT 675

FT 676

FT 677

FT 678

FT 679

FT 680

FT 681

FT 682

FT 683

FT 684

FT 685

FT 686

FT 687

FT 688

FT 689

FT 690

FT 691

FT 692

FT 693

FT 694

FT 695

FT 696

FT 697

FT 698

FT 699

FT 700

FT 701

FT 702

FT 703

FT 704

FT 705

FT 706

FT 707

FT 708

FT 709

FT 710

FT 711

FT 712

FT 713

FT 714

FT 715

FT 716

FT 717

FT 718

FT 719

FT 720

FT 721

FT 722

FT 723

FT 724

FT 725

FT 726

FT 727

FT 728

FT 729

FT 730

FT 731

FT 732

FT 733

FT 734

FT 735

FT 736

FT 737

FT 738

FT 739

FT 740

FT 741

FT 742

FT 743

FT 744

FT 745

FT 746

FT 747

FT 748

FT 749

FT 750

FT 751

FT 752

FT 753

FT 754

FT 755

FT 756

FT 757

FT 758

FT 759

FT 760

FT 761

FT 762

FT 763

FT 764

FT 765

FT 766

FT 767

FT 768

FT 769

FT 770

FT 771

FT 772

FT 773

FT 774

FT 775

FT 776

FT 777

FT 778

FT 779

FT 780

FT 781

FT 782

FT 783

FT 784

FT 785

FT 786

FT 787

FT 788

FT 789

FT 790

FT 791

FT 792

FT 793

FT 794

FT 795

FT 796

FT 797

FT 798

FT 799

FT 800

FT 801

FT 802

FT 803

FT 804

FT 805

FT 806

FT 807

FT 808

FT 809

FT 810

FT 811

FT 812

FT 813

FT 814

FT 815

FT 816

FT 817

FT 818

FT 819

FT 820

FT 821

FT 822

FT 823

FT 824

FT 825

FT 826

FT 827

FT 828

FT 829

FT 830

FT 831

FT 832

FT 833

FT 834

FT 835

FT 836

FT 837

FT 838

FT 839

FT 840

FT 841

FT 842

FT 843

FT 844

FT 845

FT 846

FT 847

FT 848

FT 849

FT 850

FT 851

FT 852

FT 853

FT 854

FT 855

FT 856

FT 857

FT 858

FT 859

FT 860

FT 861

FT 862

FT 863

FT 864

FT 865

FT 866

FT 867

FT 868

FT 869

FT 870

FT 871

FT 872

FT 873

FT 874

FT 875

FT 876

FT 877

FT 878

FT 879

FT 880

FT 881

FT 882

FT 883

FT 884

FT 885

FT 886

FT 887

FT 888

FT 889

FT 890

FT 891

FT 892

FT 893

FT 894

FT 895

FT 896

FT 897

FT 898

FT 899

FT 900

FT 901

FT 902

FT 903

FT 904

FT 905

FT 906

FT 907

FT 908

FT 909

FT 910

FT 911

FT 912

FT 913

FT 914

FT 915

FT 916

FT 917

FT 918

FT 919

FT 920

FT 921

FT 922

FT 923

FT 924

FT 925

FT 926

FT 927

FT 928

FT 929

FT 930

FT 931

FT 932

FT 933

FT 934

FT 935

FT 936

FT 937

FT 938

FT 939

FT 940

FT 941

FT 942

FT 943

FT 944

FT 945

FT 946

FT 947

FT 948

FT 949

FT 950

FT 951

FT 952

FT 953

FT 954

FT 955

FT 956

FT 957

FT 958

FT 959

FT 960

FT 961

FT 962

FT 963

FT 964

FT 965

FT 966

FT 967

FT 968

FT 969

FT 970

FT 971

FT 972

FT 973

FT 974

FT 975

FT 976

FT 977

FT 978

FT 979

FT 980

FT 981

FT 982

FT 983

FT 984

FT 985

FT 986

FT 987

FT 988

FT 989

FT 990

FT 991

FT 992

FT 993

FT 994

FT 995

FT 996

FT 997

FT 998

FT 999

FT 1000

FT 1001

FT 1002

FT 1003

FT 1004

FT 1005

FT 1006

FT 1007

FT 1008

FT 1009

FT 1010

FT 1011

FT 1012

FT 1013

FT 101

```

116 D-ANEC--MANNGGCEHCVN-TIGTY--
212 AVQGHWAREAFGAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAALQADGRSCTA----- 264
146 FELSGDNTCSDDIDECAVSNGGSCDRVNSPGGFRCDPFDLYLHADGRTCGSGFHFENL 205
265 ---SATQSC---NDLCBHFCVNPDPQSGYSVCMETGYRLAADQHRCEVDVDCILLEPSPC 318
206 ILIKVKTSCSTDNNGGCEHC-ENDSNGEFYRCRCRVGFKLSENKRSQCPVDFCNKGC 264
319 PORCVNTGGFECHCYPNYL-VDGECVEPVDPCFRAN-CETOCPPLNOTSYLVCVCAEGF 376
265 QHHCNTNNHGRAQCQCYPGFHLSDYDRSCVDIDECANKNGCEHFENVKGT-YRCKRGGY 323
377 APIPEPHRCQMF-----CNQACAPAC--DPNTQASCECPGEGYIL-DGFTCTDIDEC- 427
324 -QLGRDGRGTCEMLGGCGQVNGGCGHCDYQDPGGHVCKCRNGYIILANDQKLCHDINECH 382
428 ENGFGCSGVCHNLPGTTECIGDPSALVRHGT-----DCDS-----GKVDGGDSG 473
383 ENNGDCSQICVNLAGSVECCQPGFRLWKORKCEDISECSSNNGCGEQLCSNQEGGYMC 442
474 SCEP 477
443 SCEP 446

RESULT 12
O88281 PRELIMINARY; PRT; 1574 AA.
ID O88281 AC O88281;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MEGF6.
GN MEGF6.
NC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RT like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR ENML; AB011532; BAA32462.1; -.
DR PIR; T13954; T13954.
DR HSSP; P00736; LAPQ.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000152; Asx hydroxyl_s.
DR InterPro; IPR001881; EGF_Ca
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 20.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00179; EGF_Ca; 4.
DR PROSITE; PS00010; ASX HYDROXYL; 5.
DR PROSITE; PS00022; EGF_1; 23.
DR PROSITE; PS01186; EGF_2; 23.
DR PROSITE; PS01187; EGF_CA; 5.
DR EGF-like domain.
SQ SEQUENCE 1574 AA; 165445 MW; 2B48533DBF77F6E7 CRC64;

Query Match 14.6%; Score 414; DB 11; Length 1574;
Best Local Similarity 34.1%; Pred.No. 1.8e-25;
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 1
QY 206 CTAPPG-AVOGHWAREAFGAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAALQADGRSCTA

```

147 CRCPGGYQLQGD--GKTCQDVDECEAHNGGQCHRCVNTPGSYLCECKPGFRLHDTGRTCLD- 200

265 SATQSC---NDLQBHFQVP-----NPDQPGSYSCM-----291

205 -AISCTILNGGCGHQHCVQVLTVTHQRCRCPQYQLQEDGRRVRRPCAEAGNGGCMHI CQ 263

292 -----CBTGYRLAADQHRCEVDVDCILEFSPQRCVNTQGGFCEHCYFNVDL-VD 341

264 ELRGLAHCGCHPGYQLAADRKTTEDVDECALGAQCAHGLNTQGSFKVCVCHAGYELGAD 323

342 G-EC-----VEVPDPCPRAN--CEYQCOPLNQTSYLVCBAEGFAPIPHEPHRCQMF-----390

324 GRQCVRIEMEIVNGCEAGNGGCGSHGCSH-TSTGP LTCPRGY-----ELDEDQKTCIDID 377

391 ---NOTACPADCDNTQA--SCECPGVIID-DGFICTIDCENG-GFCSGGVCHNLPGT 443

378 DCANSPCCQQAQ-ANTPGYECSCFAGYRLNTDCCGEDVDECAHGCGGHEHCNSLAGS 436

444 PECIC 448

437 FQCF 441

RESULT 13

Q9VRS89

ID ID Q9VRS89 PRELIMINARY; PRT; 1409 AA.

AC Q9VRS89;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE CG7526 protein.

DE CG7526.

GN Drosophila melanogaster (Fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_Taxid=7227;

UN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RC MEDLINE=20196006; PubMed=107311132;

RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides F.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Bavendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J.J., Brockstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.J.,



RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banjon J., An H., Baldwin D., Banjon J., Beeson K.Y., Buesam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacieb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]

RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AEO03558; F: AEO5038.2;  
 DR HSSP: P00736; LAPO.  
 DR FlyBase: FBgn0035798; CG7526.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005509; F:calcium ion binding; IEA.  
 DR GO: GO:0004888; F:transmembrane receptor activity; IEA.  
 DR GO: GO:0007596; P:blood coagulation; IEA.  
 DR InterPro: IPR000152; Asx hydroxylase.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR003410; Hyalin.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001491; Thrombomodulin.  
 DR Pfam: PF00008; EGF; 11.  
 DR Pfam: PF00084; sushi; 2.  
 DR PRINTS: PR00907; THROMBOMODULN.  
 DR SMART: SM00032; CCP; 2.  
 DR SMART: SM00181; EGF; 13.  
 DR SMART: SM00179; EGF\_Ca; 14.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 8.  
 DR PROSITE: PS01186; EGF\_2; 9.  
 DR PROSITE: PS01187; EGF\_Ca; 10.  
 DR EGF-like domain.  
 SQ SEQUENCE 1409 AA; 154131 MW; F9A0DE1478AF354C CRC64;

Query Match 13.6%; Score 385; DB 5; Length 1409;  
 Best Local Similarity 32.1%; Pred. No. 4e-23;  
 Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;

QY 206 CTAPPGAVQG-----HWAREAPGAW-----DCSVNGGCGEACNAIPG 243

Db 563 CLCPGGYALGLDNHIVTSLSNGSFITDSTSTSPSAHTCLDIDECSLANGNSHFQCNBPG 622  
 QY 244 APRCCPAGALQADGRCTASATQSC---NDLCEHFVCPNPDPGSGSCMCETGVRLAA 300  
 Db 623 GFQACAPLGYALSEDMRTC--QDIDECUDSNGQCSQLCL---NQPGGFACACETGFELPP 677  
 QY 301 DOHRCEDVDDCILEPSPCPQRCVNTQGGFECHCYPNYDLVDGE--CVEFVDPFC---FRAN 355  
 Db 678 DDFGADIDECSDYNGNSDICIINLLGTHACACRGYELAKDKLSCLD-VDECAGLLSG 736  
 QY 356 CEYQCPINQOT-SYLCVCAEGFAPIPHPHRCQMFNCOTACPA---DCDPNTQASC----- 407  
 Db 737 CSHEC--INKAGTFECGCPGAG--TLNDDGR-----SCSPALVGCPPGTORSADGCAP 785  
 QY 408 -PCPGYILDDGFICTDIDEC--ENGSCFSGVCHNLPGTFECICGPF-----DSALVRHI 458  
 Db 786 IECNFGYTLGSDKCDVIDECQKQNGG-CSHRCSNTEGSKFCSCPPGYELDSQDKTCQDI 844  
 QY 459 GTDCDSGK 466  
 Db 845 -DECQDQK 851

## RESULT 14

P87363 PRELIMINARY; PRT; 708 AA.  
 AC P87363;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Fibrillin-1 (Fragment).  
 GN FBNI.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20152896; PubMed=10691037;  
 RA Zhou G., Price C.E., Rosenquist T.H., Gadsen P.F., Godfrey M.;  
 RT "Partial cloning and sequencing of chick fibrillin-1 cDNA.";  
 RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).  
 DR EMBL: U88872; AAB48531.1; -.  
 DR HSSP: P07204; 2ADX.  
 DR GO: GO:0005509; F:calcium ion binding; IEA.  
 DR InterPro: IPR000152; Asx hydroxylase.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR002212; Fibril-assoc.  
 DR Pfam: PF00008; EGF; 13.  
 DR Pfam: PF00683; TB; 2.  
 DR SMART: SM00179; EGF\_Ca; 14.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 13.  
 DR PROSITE: PS01186; EGF\_2; 10.  
 DR PROSITE: PS01187; EGF\_Ca; 13.  
 DR EGF-like domain.  
 KW EGF-like domain.  
 FT NON TER 708  
 SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

Query Match 13.3%; Score 375; DB 13; Length 708;  
 Best Local Similarity 31.8%; Pred. No. 1.2e-22;  
 Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

QY 226 DCSVNGGCGEACNAIPGAPRCQCPAGALQADGRCTASATQSCNDLCEHFVCPNP--- 282  
 Db 223 ECSIMNGCGENFCTSGSEYECSCQGFALMPDHRCTCT-----DIDB--CEDNPNIC 272  
 QY 283 -----DQPGSYSCMCETGVRLAADHRCEDVDDCILEPSPC-PQRCVNTQGGFECHCY 335  
 Db 273 DGGQCTNIPGEYRCVLCVDGFMASDMKTCVDVNECDLHFNICLSGTCENTKGSFICHCDM 332



```
QY 336 NYDLVDGE--CVPSPVDP--FRANCEYQCQPLN--QTSYLCVCAEGFA-----PIP 380
Db 333 GYSKXKTTCTD--INECEIGAHCNDRHAVCTNIPGSFKSCSSGMIGNGIKCTDLDECS 391
QY 381 HEPHRCQMFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDE-----CENG- 430
Db 392 NGTHKCSPH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDCESENINLCENGQ 442
QY 431 -----GF-----CS-----GVCHNLPGTFECICGPDPSALV 455
Db 443 CLNAPGGYRCEDMGFLPSLDGKACEDIDECSLPNICVYGTCNHLPGLFRCCEBVGVELD 502
QY 456 RHIG-----TPCDGSK-VDGSQSGSEPPSPFTPGSTLTTPAVGLV 495
Db 503 RSGNCTDVNECAPTTCTISGTCVNTAGSYTCECPDP-----FELNPTRVGCV 550
```

```
RESULT 15
Q9WUH9 PRELIMINARY; PRT; 2906 AA.
AC Q9WUH9;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fibrillin-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99350231; PubMed=10419698;
RA Yang O., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E.,
RA Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
RL morphogenesis of embryonic lung."
RL Dev. Biol. 212:229-242 (1999).
DR EMBL; AF135060; AAD34439.1; -.
DR HSSP; P35555; 1EMN.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; A:aldehyde dehydr.
DR InterPro; IPR00152; Asx hydroxyl_S.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 44.
DR Pfam; PF00683; TB; 9.
DR SMART; SM00179; EGF_CA; 42.
DR PROSITE; PS00070; ALDEHYDE DEHYDR_CYS; 1.
DR PROSITE; PS00010; ASX HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW EGF-like domain.
SQ SEQUENCE 2906 AA; 313371 MW; 9EB64E727044EF58 CRC64;
Query Match 13.2%; Score 374; DB 11; Length 2906;
Best Local Similarity 32.9%; Pred. No. 7.3e-22;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;
QY 226 DCSVENGCEHACNAIPGAPCQCPAGALQADGRSCTASATQSCNLDCEHFCVNPED-- 283
Db 1237 ECTMNGGCDTQCTNSGSEYSCSGSEGYALMPDGRSCA-----DIDE--CENNPDIC 1286
QY 284 -----QPGSYSCWCEGYRLADQHRCEVDVDCILLEPSPCP--QRCVNTQGGFECHCYP 335
Db 1287 DGGQCTNIPGEYRCLCYDGFNASMDMTCTIDVNECDLNPNCMPGECENTKGSFICHQQL 1346
QY 336 NYDLVDGE--CVPSPVDP--FRANCEYQCQPLN--QTSYLCVCAEGFA-----PIP 380
Db 1347 GYSVKKGATGCTD--VDECEIGAHCNDRHAVCTNIPGSFKSCSSGMIGNGIKCTDLDECA 1405
```

```
QY 381 HEPHRCQMFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDE-----CENG- 430
Db 1406 NGTHQCSI-----NAQC-VNTPGYSYRCACSEGF-TGDGFTCSVDYDECAENINLCENGQ 1456
QY 431 -----GF-----CS-----GVCHNLPGTFECICGPDPSALV 455
Db 1457 CLNVPGAYRCEBWMGFTPASDSRSCQDIDECSPONICVEFTCNLPGMFHCICDDGYGLD 1516
QY 456 RHIG--TDGD 463
Db 1517 RTGGHCTDID 1526
Search completed: June 9, 2004, 08:49:59
Job time : 40.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: June 9, 2004, 08:39:53 ; Search time 53.5 Seconds  
(without alignments)  
2630.068 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516  
Perfect score: 2830  
Sequence: 1 APAEPQGGSCVCHDFAL.....PSPTGSLTPPAVLGHVG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003s:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID         | Description        |
|------------|-------|-------|--------|------------|--------------------|
| 1          | 2830  | 100.0 | 516    | 2 AAY09348 | Aay09348 Human thr |
| 2          | 2830  | 100.0 | 516    | 3 AAY83935 | Aay83935 Human thr |
| 3          | 2830  | 100.0 | 516    | 6 ABU08703 | Abu08703 Thrombomo |
| 4          | 2830  | 100.0 | 516    | 7 ADA37275 | Ada37275 Human sol |
| 5          | 2830  | 100.0 | 575    | 6 ABG72575 | Abg72575 Human thr |
| 6          | 2830  | 100.0 | 575    | 7 ABM78950 | Abm78950 Breast ca |
| 7          | 2830  | 100.0 | 575    | 7 ADE54760 | Ades4760 Human Pro |
| 8          | 2830  | 100.0 | 575    | 7 ADD48168 | Add48168 Human Pro |
| 9          | 2830  | 100.0 | 631    | 6 ABG72572 | Abg72572 Human thr |
| 10         | 2826  | 99.9  | 498    | 2 AAR84185 | Aar84185 Human der |
| 11         | 2826  | 99.9  | 498    | 2 AAW01600 | Aaw01600 Thrombomo |
| 12         | 2826  | 99.9  | 498    | 2 AAR22016 | Aar22016 Truncated |
| 13         | 2826  | 99.9  | 516    | 2 AAY09347 | Aay09347 Human thr |
| 14         | 2826  | 99.9  | 516    | 2 AAY83934 | Aay83934 Human thr |
| 15         | 2826  | 99.9  | 516    | 3 AAY69529 | Aay69529 Human thr |
| 16         | 2826  | 99.9  | 516    | 3 ABU08701 | Abu08701 Thrombomo |
| 17         | 2826  | 99.9  | 516    | 7 ADA37271 | Ada37271 Human sol |
| 18         | 2826  | 99.9  | 516    | 7 AAR11534 | Aar11534 Human thr |
| 19         | 2826  | 99.9  | 575    | 2 AAR11806 | Aar11806 Thrombomo |
| 20         | 2826  | 99.9  | 575    | 2 AAR41031 | Aar41031 Human thr |
| 21         | 2826  | 99.9  | 575    | 5 AAE17521 | Aae17521 Human ful |
| 22         | 2826  | 99.9  | 575    | 5 AAE23026 | Aae23026 Human thr |
| 23         | 2826  | 99.9  | 575    | 5 AAR94607 | Aar94607 Human rec |
| 24         | 2824  | 99.8  | 497    | 2 AAR10617 | Aar10617 Soluble t |
| 25         | 2824  | 99.8  | 515    | 2 AAR10617 | Aar10617 Soluble t |

|    |      |      |     |            |                    |
|----|------|------|-----|------------|--------------------|
| 26 | 2824 | 99.8 | 575 | 1 AAP82070 | Aap82070 Human thr |
| 27 | 2823 | 99.8 | 557 | 5 AAE17528 | Aae17528 Human thr |
| 28 | 2823 | 99.8 | 557 | 5 AAE17525 | Aae17525 Human thr |
| 29 | 2823 | 99.8 | 557 | 5 AAE23031 | Aae23031 Human thr |
| 30 | 2823 | 99.8 | 557 | 5 AAE23028 | Aae23028 Human thr |
| 31 | 2822 | 99.7 | 516 | 2 AAR22018 | Aar22018 Human thr |
| 32 | 2822 | 99.7 | 557 | 5 AAE17523 | Aae17523 Human thr |
| 33 | 2822 | 99.7 | 575 | 2 AAR31572 | Aar31572 Human thr |
| 34 | 2821 | 99.7 | 557 | 5 AAE17596 | Aae17596 Human thr |
| 35 | 2820 | 99.6 | 497 | 2 AAW69520 | Aaw69520 rSTM prot |
| 36 | 2820 | 99.6 | 557 | 5 AAE17530 | Aae17530 Human thr |
| 37 | 2820 | 99.6 | 557 | 5 AAE17526 | Aae17526 Human thr |
| 38 | 2819 | 99.6 | 557 | 5 AAE17529 | Aae17529 Human thr |
| 39 | 2819 | 99.6 | 557 | 5 AAE23029 | Aae23029 Human thr |
| 40 | 2819 | 99.6 | 575 | 6 ABU52408 | Abu52408 Human GPC |
| 41 | 2818 | 99.6 | 516 | 2 AAR22017 | Aar22017 Human thr |
| 42 | 2818 | 99.6 | 557 | 5 AAE17593 | Aae17593 Human thr |
| 43 | 2818 | 99.6 | 557 | 5 AAE17595 | Aae17595 Human thr |
| 44 | 2818 | 99.6 | 557 | 5 AAE17597 | Aae17597 Human thr |
| 45 | 2818 | 99.6 | 557 | 5 AAE17529 | Aae17529 Human thr |

ALIGNMENTS

RESULT 1  
AAY09348  
ID AAY09348 standard; protein; 516 AA.  
XX  
AC AAY09348;  
XX  
XX  
DT 08-JUL-1999 (first entry)  
XX  
DE Human thrombomodulin SEQ ID NO:2.  
XX  
KW Human; thrombomodulin; aqueous parenteral solution; storage;  
KW distribution; acute coronary syndrome; thrombosis; embolism; diabetes.  
XX  
OS Homo sapiens.  
XX  
FN WO9918994-A1.  
XX  
XX  
PD 22-APR-1999.  
XX  
PF 13-OCT-1998; 98WO-JP004603.  
XX  
PR 15-OCT-1997; 97JP-00281659.  
PR 11-NOV-1997; 97JP-00308523.  
XX  
(ASAH ) ASAH KASEI KOGYO KK.  
XX  
PI Yui M, Yokozawa A, Murata T, Tsuruta K, Shimizu H;  
XX  
DR WPI; 1999-277444/23.  
XX  
N-PSDB; AAX55880.  
XX  
PT Stable aqueous parenteral thrombomodulin solution - comprising buffer and  
PT surfactant, useful for treating acute coronary syndrome, thrombosis,  
PT embolism, and diabetes.  
XX  
PS Claim 7; Page 90-92; 97pp; Japanese.  
XX

The present invention describes a method for maintaining the quality of an aqueous parenteral solution of thrombomodulin comprising buffer and surfactant aseptically filled in a case or syringe. Maintaining the quality of an aqueous, parenteral thrombomodulin solution is characterised in that the solution: (a) comprises soluble thrombomodulin, a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically filled into: (i) a case or (ii) a syringe without any empty space; and (c) is kept in liquid form in storage and distribution and not frozen or freeze dried. Thrombomodulin is used to treat acute coronary syndrome (such as myocardial infarction, unstable angina and coronary artery blockage), thrombosis (e.g. cerebral, vascular and peripheral blood



Wed Jun 9 13:16:02 2004

XX OS Unidentified.  
 XX PN WO2003013606-A1.  
 XX PD 20-FEB-2003.  
 XX PF 07-AUG-2002; 2002WO-JP008058.  
 XX PR 08-AUG-2001; 2001JP-00241371.  
 XX PA (ASAH) ASAHI KASEI KOGYO KK.  
 XX PI Suzuki H, Mohri M;  
 XX DR WPI; 2003-248241/24.  
 XX DR N-PSDB; ABX94089.  
 XX PT Thrombomodulin-containing medicinal compositions useful for treating or  
 PT preventing e.g. myocardial infarction, brain infarction, acute coronary  
 PT syndrome and sepsis.  
 XX PS Claim 18; Page 34-37; 48pp; Japanese.  
 XX CC The invention describes drug compositions containing thrombomodulin, and  
 CC an antiplatelet, an anticoagulant or/ and a thrombolytic other than  
 CC thrombomodulin. The drug compositions are for preventing or treating  
 CC diseases or syndromes due to abnormality in the blood coagulation-  
 CC fibrinolysis system or/ and vascular obstruction, including myocardial  
 CC infarction, acute coronary syndrome, brain infarction and sepsis,  
 CC particularly as combination drug compositions. The compositions have high  
 CC therapeutic efficacy. This is the amino acid sequence of a thrombomodulin  
 CC -containing composition associated protein  
 XX SQ Sequence 516 AA;  
 Query Match 100.0%; Score 2830; DB 6; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-150;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVISLLNGD 60  
 DB 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVISLLNGD 78  
 QY 61 GGVGRRRLWTGLQPPCGGPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
 DB 79 GGVGRRRLWTGLQPPCGGPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEQCEVKADGFLCEPHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 DB 139 AVSAAEATVPSEPIWEQCEVKADGFLCEPHFPATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAPGAWDCSVENGCGEHACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAPGAWDCSVENGCGEHACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATOSCNLCEHFCVPNPDPQPGSYSCMCETGYRLAA 300  
 DB 259 IPGAPRCQCPAGAAALQADGRSCTASATOSCNLCEHFCVPNPDPQPGSYSCMCETGYRLAA 318  
 QY 301 DQHRCEVDVDDCILEPSPCPCQRCVNTQGGFECHECHYCNPDYLDVDCGCVPEVDPFCFRANCEYQC 360  
 DB 319 DQHRCEVDVDDCILEPSPCPCQRCVNTQGGFECHECHYCNPDYLDVDCGCVPEVDPFCFRANCEYQC 378  
 QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPNTQASCECEGYLLDDGFI 420  
 DB 379 QPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQTACPADCDPNTQASCECEGYLLDDGFI 438  
 QY 421 CTTDIDCEGNGGFCGVCNHLPGTFECICGPDLSALARIHIGTDCDSGKVDGSGSGGPPPS 480  
 DB 439 CTTDIDCEGNGGFCGVCNHLPGTFECICGPDLSALARIHIGTDCDSGKVDGSGSGGPPPS 498  
 QY 481 PTFGSTLTTPPAVGLVHSG 498

DB 499 PTFGSTLTTPPAVGLVHSG 516  
 RESULT 4  
 ADA37275  
 ID ADA37275 standard; protein; 516 AA.  
 XX AC ADA37275;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Human soluble thrombomodulin protein SEQ ID NO:5.  
 XX KW high-concentration preparation; soluble thrombomodulin; human;  
 XX KW thrombomodulin.  
 XX OS Homo sapiens.  
 XX PN WO2003061687-A1.  
 XX PD 31-JUL-2003.  
 XX PF 17-JAN-2003; 2003WO-JP000339.  
 XX PR 18-JAN-2002; 2002JP-00009951.  
 XX PA (ASAH) ASAHI KASEI KK.  
 XX PI Nishio F;  
 XX DR WPI; 2003-712487/67.  
 XX DR N-PSDB; ADA37276.  
 XX PT High concentration stable low-foaming soluble thrombomodulin preparation  
 PT for pharmaceutical use.  
 XX PS Claim 18; Page 102-104; 113pp; Japanese.  
 XX CC The present invention describes the high-concentration preparation of  
 CC soluble thrombomodulin. The present sequence represents a human soluble  
 CC thrombomodulin, which is given in the exemplification of the present  
 CC invention.  
 XX SQ Sequence 516 AA;  
 Query Match 100.0%; Score 2830; DB 7; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-150;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVISLLNGD 60  
 DB 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVISLLNGD 78  
 QY 61 GGVGRRRLWTGLQPPCGGPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
 DB 79 GGVGRRRLWTGLQPPCGGPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEQCEVKADGFLCEPHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 DB 139 AVSAAEATVPSEPIWEQCEVKADGFLCEPHFPATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAPGAWDCSVENGCGEHACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAPGAWDCSVENGCGEHACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATOSCNLCEHFCVPNPDPQPGSYSCMCETGYRLAA 300  
 DB 259 IPGAPRCQCPAGAAALQADGRSCTASATOSCNLCEHFCVPNPDPQPGSYSCMCETGYRLAA 318  
 QY 301 DQHRCEVDVDDCILEPSPCPCQRCVNTQGGFECHECHYCNPDYLDVDCGCVPEVDPFCFRANCEYQC 360  
 DB 319 DQHRCEVDVDDCILEPSPCPCQRCVNTQGGFECHECHYCNPDYLDVDCGCVPEVDPFCFRANCEYQC 378

|          |             |   |     |
|----------|-------------|---|-----|
| QY       | 361         | QPLNQTSLVCAEGFAPIPHEPHRCQMFQNTACPADCDNTQASCECPGSIILDDGFI                  | 420 |
| DB       | 379         | QPLNQTSLVCAEGFAPIPHEPHRCQMFQNTACPADCDNTQASCECPGSIILDDGFI                  | 438 |
| QY       | 421         | CTDIDECENGFCGSGVCHNLPTFFCI  | 480 |
| DB       | 439         | CTDIDECENGFCGSGVCHNLPTFFCI  | 498 |
| QY       | 481         | PTPGSTLTTPPAVGLVHSG   | 498 |
| DB       | 499         | PTPGSTLTTPPAVGLVHSG   | 516 |
| RESULT 5 |             |   |     |
| ID       | ABG72575    | standard; protein; 575 AA.  |     |
| XX       | AC          | ABG72575;   |     |
| DT       | 05-MAR-2003 | (first entry)   |     |
| XX       | XX          | Human thrombomodulin protein #2 (CD141).                                  |     |
| XX       | XX          | Human; thrombomodulin; CD141; dendritic cell separation; inflammation;    |     |
| XX       | XX          | anti-inflammatory; cytostatic; antiarthritic; antirheumatic;              |     |
| XX       | XX          | immunosuppressive; dermatological; cell therapy; dermatomyositis; cancer; |     |
| XX       | XX          | autoimmune disorder; rheumatoid arthritis; psoriatic arthritis;           |     |
| XX       | XX          | alkalizing spondylitis; Sjorgen's syndrome; lupus erythematosus;          |     |
| XX       | XX          | Goodpasture's syndrome; scleroderma; vasculitis; polymyositis.            |     |
| XX       | OS          | Homo sapiens.   |     |
| XX       | XX          | WO200293172-A1.   |     |
| XX       | XX          | 21-NOV-2002.  |     |
| XX       | XX          | 17-MAY-2002; 2002WO-US015786.   |     |
| XX       | XX          | 17-MAY-2001; 2001US-0291561P.   |     |
| XX       | XX          | (MILT-) MILTENYI BIOTEC INC.  |     |
| XX       | XX          | Schmitz J, Dziunek A, Buck DW;  |     |
| XX       | XX          | WPI; 2003-129324/12.  |     |
| XX       | XX          | Enriching for dendritic cells from a mixture of cells, useful for         |     |
| XX       | XX          | treating inflammation, cancer or autoimmune disorders such as arthritis,  |     |
| XX       | XX          | or lupus by contacting the mixture of cell with an antigen-binding        |     |
| XX       | XX          | fragment specific for CD141.  |     |
| XX       | XX          | Example 1; Fig 3B; 88pp; English.   |     |
| XX       | XX          | This invention relates to a novel method for separating dendritic cells   |     |
| XX       | XX          | from a mixture of cells. The method comprises contacting the mixture of   |     |
| XX       | XX          | cells with an antigen-binding fragment specific for CD141, and selecting  |     |
| XX       | XX          | the cells that are CD141 positive therefore producing a dendritic cell-   |     |
| XX       | XX          | enriched composition. The method of the invention may have                |     |
| XX       | XX          | antiinflammatory; cytostatic; antiarthritic; antirheumatic;               |     |
| XX       | XX          | immunosuppressive and dermatological activities and may be used in cell   |     |
| XX       | XX          | therapy. The methods, compositions and agents of the invention are useful |     |
| XX       | XX          | for treating inflammation, cancer or autoimmune disorders such as         |     |
| XX       | XX          | rheumatoid arthritis, psoriatic arthritis, alkalizing spondylitis,        |     |
| XX       | XX          | Sjorgen's syndrome, lupus erythematosus, Goodpasture's syndrome,          |     |
| XX       | XX          | scleroderma, vasculitis, polymyositis or dermatomyositis. The dendritic   |     |
| XX       | XX          | cells and populations are useful for screening assays for agents that     |     |
| XX       | XX          | affect dendritic cell function, and for the phenotypic, functional,       |     |
| XX       | XX          | biochemical or molecular analyses of dendritic cells. The compositions    |     |
| XX       | XX          | and dendritic cells are useful as vaccine adjuvants. The present sequence |     |
| XX       | XX          | represents a protein sequence of human CD141 protein which is used to     |     |
| XX       | XX          | generate the dendritic cell specific antibodies that are used in the      |     |
| XX       | XX          | method of the invention   |     |
| XX       | SQ          | Sequence 575 AA;  |     |
| XX       | QY          | Query Match   |     |
| XX       | DB          | Best Local Similarity 100.0%; Score 2830; DB 6; Length 575;               |     |
| XX       | QY          | Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;              |     |
| XX       | DB          | 1 APASPOPGSGOCVHDCFALYFGPATFLNASQICDGLRGHLMTVRSSVAADVISILLNGD 60          |     |
| XX       | DB          | 19 APAEPQPGSGOCVHDCFALYFGPATFLNASQICDGLRGHLMTVRSSVAADVISILLNGD 78         |     |
| XX       | QY          | 61 GGVGRRRLWIGLQPLPPGCGDKRIGLPIRGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120       |     |
| XX       | DB          | 79 GGVGRRRLWIGLQPLPPGCGDKRIGLPIRGFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 138       |     |
| XX       | QY          | 121 AVSAAEATVPSEPIWEEQOCEVKADGFLCEFHFPATCRPLAVEPGAAAAAASITYGTPEA 180      |     |
| XX       | DB          | 139 AVSAAEATVPSEPIWEEQOCEVKADGFLCEFHFPATCRPLAVEPGAAAAAASITYGTPEA 198      |     |
| XX       | QY          | 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 240     |     |
| XX       | DB          | 199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAPGAWDCSVENGCGCEHACNA 258     |     |
| XX       | QY          | 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300        |     |
| XX       | DB          | 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318        |     |
| XX       | QY          | 301 DOHRCEVDVDCILPEPSPQPCQPCVNTQGGFECHCYPNYDLVDGECVEVPDFCFRANCEYQC 360    |     |
| XX       | DB          | 319 DOHRCEVDVDCILPEPSPQPCQPCVNTQGGFECHCYPNYDLVDGECVEVPDFCFRANCEYQC 378    |     |
| XX       | QY          | 361 QPLNQTSLVCAEGFAPIPHEPHRCQMFQNTACPADCDNTQASCECPGSIILDDGFI 420          |     |
| XX       | DB          | 379 QPLNQTSLVCAEGFAPIPHEPHRCQMFQNTACPADCDNTQASCECPGSIILDDGFI 438          |     |
| XX       | QY          | 421 CTDIDECENGFCGSGVCHNLPTFFCI  | 480 |
| XX       | DB          | 439 CTDIDECENGFCGSGVCHNLPTFFCI  | 498 |
| XX       | QY          | 481 PTPGSTLTTPPAVGLVHSG   | 498 |
| XX       | DB          | 499 PTPGSTLTTPPAVGLVHSG   | 516 |
| RESULT 6 |             |   |     |
| XX       | ABM78950    | standard; protein; 575 AA.  |     |
| XX       | AC          | ABM78950;   |     |
| XX       | AC          | ABM78950;   |     |
| XX       | DT          | 15-JAN-2004 (first entry)   |     |
| XX       | XX          | Breast cancer specific marker under-expressed in breast cancer.           |     |
| XX       | XX          | Breast cancer; marker; thrombomodulin; human; diagnosis; cytostatic;      |     |
| XX       | XX          | biochip; vaccine.   |     |
| XX       | OS          | Homo sapiens.   |     |
| XX       | XX          | WO2003073911-A2.  |     |
| XX       | XX          | 12-SEP-2003.  |     |
| XX       | XX          | 27-FEB-2003; 2003WO-US005984.   |     |
| XX       | XX          | 28-FEB-2002; 2002US-0359999P.   |     |
| XX       | XX          | (GEO ) UNIV GEORGETOWN.   |     |
| XX       | XX          | Su Ya, Yang J;  |     |
| XX       | XX          | WPI; 2003-721995/68.  |     |
| XX       | XX          | N-PSDB; ACP79930.   |     |

XX Detecting breast cancer in a subject comprises contacting a biological  
 PT sample with an agent that binds to a polynucleotide or polypeptide of a  
 TT breast-cancer specific gene (BCSG).  
 XX  
 XX Claim 5; Page 119-121; 143pp; English.  
 XX  
 XX The present sequence is that of a breast cancer specific marker (BCSM)  
 CC encoded by a gene identified by microarray gene expression analysis as  
 CC being under-expressed in breast cancers in comparison to healthy tissue.  
 CC The BCSM was identified as thrombomodulin (THBD), a negative regulator of  
 CC coagulation previously reported to be involved in vascular diseases and  
 CC cancers. It is 1 of 19 (see ABW78941-59) BCSMs of the invention that are  
 CC encoded by breast cancer specific genes (BCSGs) which are differentially  
 CC expressed in breast cancer cell lines and breast cancer tissue samples as  
 CC compared to control cell lines and normal tissue samples. The invention  
 CC provides a method for detection of breast cancer by measuring expression  
 CC levels of BCSGs, and in particular the level of polynucleotides  
 CC transcribed from and polypeptides encoded by the BCSGs. A pharmaceutical  
 CC composition for the treatment of breast cancer comprises a BCSM, an  
 CC antibody directed against a BCSM, a vaccine generated using a BCSM, or an  
 CC agent that modulates an expression level of a BCSG or an activity of a  
 CC BCSM. A biochip for diagnosing breast cancer or screening agents that  
 CC inhibit breast cancer comprises a BCSG or BCSM  
 XX Sequence 575 AA;

Query Match 100.0%; Score 2830; DB 7; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-150;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APAEPQGGSCQVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFWQVTGDNNTSYSRWARLDINGAPLCGLCV 120  
 Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFWQVTGDNNTSYSRWARLDINGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEQCEVADGFLCEFHFPATCRPLAVEPGHAAAVSITVTGTPFA 180  
 Db 139 AVSAAEATVPSEPIWEQCEVADGFLCEFHFPATCRPLAVEPGHAAAVSITVTGTPFA 198  
 QY 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGANDCSVENGCGEHA 240  
 Db 199 ARGADFQALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGANDCSVENGCGEHA 258  
 QY 241 IPGAPRCOCAPAGALQADGSCSTASATQSCNDICEHFCVNPDPGYSYCMCTGVRLAA 300  
 Db 259 IPGAPRCOCAPAGALQADGSCSTASATQSCNDICEHFCVNPDPGYSYCMCTGVRLAA 318  
 QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCYPNYDLVDGCEVPDPFCFRANCEYQC 360  
 Db 319 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCYPNYDLVDGCEVPDPFCFRANCEYQC 378  
 QY 361 QPLNQTSLVCARGFAPIEPHRCPMPCNQACADCPNTQASCEPGEVILDDGFI 420  
 Db 379 QPLNQTSLVCARGFAPIEPHRCPMPCNQACADCPNTQASCEPGEVILDDGFI 438  
 QY 421 CTDIDECNGGFCGVCNHLPGTTEPCICGPDALAHIGHITDCCGKVGDDSGSGEP 480  
 Db 439 CTDIDECNGGFCGVCNHLPGTTEPCICGPDALAHIGHITDCCGKVGDDSGSGEP 498  
 QY 481 PTFPGSTLTPPAVLVHSG 498  
 Db 499 PTFPGSTLTPPAVLVHSG 516

RESULT 7  
 ADE54760  
 ID ADE54760 standard; protein; 575 AA.  
 XX  
 AC ADE54760;

XX 29-JAN-2004 (first entry)  
 XX Human Protein P07204, SEQ ID NO 565.  
 DE  
 XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX Homo sapiens.  
 OS  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 XX 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX  
 XX Woolf C, D'urso D, Befort K, Costigan M;  
 PI MPI; 2003-268312/26.  
 XX GENBANK; P07204.  
 DR  
 XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 XX Claim 1; Page; 1017pp; English.  
 XX  
 CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC the sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/publiished\_pct\_sequences.  
 XX  
 SQ Sequence 575 AA;

Query Match 100.0%; Score 2830; DB 7; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-150;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 APAEPQGGSCQVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFWQVTGDNNTSYSRWARLDINGAPLCGLCV 120

derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which differentially expresses in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Query Match 100.0%; Score 2830; DB 7; Length 575;  
Best Local Similarity 100.0%; Pred. No. 8.3e-150;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78  
61 GGVGRRRLWIGLQPLPPGGDPKRLGRLGFQWVTGDNNTSYSRWARLDINGAPLGLCV 120  
79 GGVGRRRLWIGLQPLPPGGDPKRLGRLGFQWVTGDNNTSYSRWARLDINGAPLGLCV 138  
121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAAVSITGTPFA 180  
139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAAVSITGTPFA 198  
181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240  
199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258  
241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQGSYSCMCEGYRLAA 300  
259 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQGSYSCMCEGYRLAA 318  
301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGECVPPVDFCFRANCEYQC 360  
319 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGECVPPVDFCFRANCEYQC 378  
361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
421 CTDIDECENGFCGVCVCHNLPGTFCICGPDPSALARIHIGTDCDSKVDGSGSGEPPPS 480  
439 CTDIDECENGFCGVCVCHNLPGTFCICGPDPSALARIHIGTDCDSKVDGSGSGEPPPS 498  
481 PTPGSTLTTPPAVGLVHSG 498  
499 PTPGSTLTTPPAVGLVHSG 516

RESULT 9  
ABG72572  
ID ABG72572 standard; protein; 631 AA.  
XX

79 GGVGRRRLWIGLQPLPPGGDPKRLGRLGFQWVTGDNNTSYSRWARLDINGAPLGLCV 138  
121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAAVSITGTPFA 180  
139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPGAAAAAVSITGTPFA 198  
181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240  
199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258  
241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQGSYSCMCEGYRLAA 300  
259 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQGSYSCMCEGYRLAA 318  
301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGECVPPVDFCFRANCEYQC 360  
319 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGECVPPVDFCFRANCEYQC 378  
361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
421 CTDIDECENGFCGVCVCHNLPGTFCICGPDPSALARIHIGTDCDSKVDGSGSGEPPPS 480  
439 CTDIDECENGFCGVCVCHNLPGTFCICGPDPSALARIHIGTDCDSKVDGSGSGEPPPS 498  
481 PTPGSTLTTPPAVGLVHSG 498  
499 PTPGSTLTTPPAVGLVHSG 516

RESULT 8  
ADD48168  
ID ADD48168 standard; protein; 575 AA.  
XX  
AC ADD48168;  
XX  
29-JAN-2004 (first entry)  
XX  
Human Protein NP\_000352, SEQ ID NO 13866.  
XX  
Human; pain; neuronal tissue; gene therapy;  
XX  
spinal segmental nerve injury; chronic constriction injury; CCI;  
XX  
spared nerve injury; SNI; Chung.  
XX  
Homo sapiens.  
XX  
WO2003016475-A2.  
XX  
27-FEB-2003.  
XX  
14-AUG-2002; 2002WO-US025765.  
XX  
14-AUG-2001; 2001US-0312147P.  
XX  
01-NOV-2001; 2001US-0346382P.  
XX  
26-NOV-2001; 2001US-0333347P.  
XX  
(GEO) GEN HOSPITAL CORP.  
XX  
PA (FARB) BAYER AG.  
XX  
PI Woolf C, D'urso D, Belfort K, Costigan M;  
XX  
WPI; 2003-268312/26.  
XX  
GENBANK; NP\_000352.  
XX  
New composition comprising two or more isolated polypeptides, useful for  
XX  
preparing a medicament for treating pain in an animal.  
XX  
Claim 1; Page; 1017pp; English.  
XX  
The invention discloses a composition comprising two or more isolated rat  
XX  
or human polynucleotides or a polynucleotide which represents a fragment,  
XX  
CC

AC ABG72572;  
 XX 05-MAR-2003 (first entry)  
 DT Human thrombomodulin protein #1 (CD141).  
 DE  
 XX Human; thrombomodulin; CD141; dendritic cell separation; inflammation;  
 KW antiinflammatory; cytostatic; antiarthritic; antirheumatic;  
 KW immunosuppressive; dermatological; cell therapy; dermatomyositis; cancer;  
 KW autoimmune disorder; rheumatoid arthritis; psoriatic arthritis;  
 KW alkylating spondylitis; Sjorgen's syndrome; lupus erythematosus;  
 KW Goodpasture's syndrome; scleroderma; vasculitis; polymyositis.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200293172-A1.  
 FN  
 XX 21-NOV-2002.  
 PD  
 XX 17-MAY-2002; 2002WO-US015786.  
 PF  
 XX 17-MAY-2001; 2001US-0291561P.  
 PR  
 XX (MILT-) MILTENYI BIOTEC INC.  
 PA  
 XX Schmitz J, Dzionek A, Buck DW;  
 FI  
 XX WPI; 2003-129324/12.  
 DR  
 XX Enriching for dendritic cells from a mixture of cells, useful for  
 PT treating inflammation, cancer or autoimmune disorders such as arthritis,  
 PT or lupus by contacting the mixture of cell with an antigen-binding  
 PT fragment specific for CD141.  
 XX  
 XX Example 1; Fig 3A; 88pp; English.  
 PS  
 XX This invention relates to a novel method for separating dendritic cells  
 CC from a mixture of cells. The method comprises contacting the mixture of  
 CC cells with an antigen-binding fragment specific for CD141, and selecting  
 CC the cells that are CD141 positive therefore producing a dendritic cell-  
 CC enriched composition. The method of the invention may have  
 CC antiinflammatory; cytostatic; antiarthritic; antirheumatic;  
 CC immunosuppressive and dermatological activities and may be used in cell  
 CC therapy. The methods, compositions and agents of the invention are useful  
 CC for treating inflammation, cancer or autoimmune disorders such as  
 CC rheumatoid arthritis, psoriatic arthritis, alkylating spondylitis,  
 CC Sjorgen's syndrome, lupus erythematosus, Goodpasture's syndrome,  
 CC scleroderma, vasculitis, polymyositis or dermatomyositis. The dendritic  
 CC cells and populations are useful for screening assays for agents that  
 CC affect dendritic cell function, and for the phenotypic, functional,  
 CC biochemical or molecular analyses of dendritic cells. The compositions  
 CC and dendritic cells are useful as vaccine adjuvants. The present sequence  
 CC represents a protein sequence of human CD141 protein which is used to  
 CC generate the dendritic cell specific antibodies that are used in the  
 CC method of the invention  
 XX  
 XX Sequence 631 AA;  
 SQ

Query Match 100.0%; Score 2830; DB 6; Length 631;  
 Best Local Similarity 100.0%; Pred. No. 9e-150;  
 Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYFGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 60  
 DB 75 APAEPQGGSCQVEHDCFALYFGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 134  
 QY 61 GGVGRRRLWIGLQPLPGCGDKPKRLGRLGQWTVGDMNTYSRWABLDLNGAPLCPCLV 120  
 DB 135 GGVGRRRLWIGLQPLPGCGDKPKRLGRLGQWTVGDMNTYSRWABLDLNGAPLCPCLV 194  
 QY 121 AVSAAEATVPSEPIWEQCEVKADGFLCFEHPFATCRPLAVEPAAAAAASVITVGTTPA 180  
 DB 195 AVSAAEATVPSEPIWEQCEVKADGFLCFEHPFATCRPLAVEPAAAAAASVITVGTTPA 254

QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPGAVOGHWAREAPGAWDCSVENGCGEHACNA 240  
 DB 255 ARGADFOALPVGSSAAVAPLGLQLMCTAPGAVOGHWAREAPGAWDCSVENGCGEHACNA 314  
 QY 241 IGAERCCOCPAGALOADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCETGYELAA 300  
 DB 315 IGAERCCOCPAGALOADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCETGYELAA 374  
 QY 301 DQHRCEVDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVFPVDFCFRANCEYQC 360  
 DB 375 DQHRCEVDVDDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVFPVDFCFRANCEYQC 434  
 QY 361 QPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 420  
 DB 435 QPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 494  
 QY 421 CTDIDECENGFCGVCVCHNLPFTFECICGPDPSALAHIGTDCDSKGVGDSGSGFPFPPS 480  
 DB 495 CTDIDECENGFCGVCVCHNLPFTFECICGPDPSALAHIGTDCDSKGVGDSGSGFPFPPS 554  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 555 PTPGSTLTTPPAVGLVHSG 572

RESULT 10  
 AAR84185  
 ID AAR84185 standard; protein; 498 AA.  
 AC AAR84185;  
 XX  
 DT 05-JUN-1996 (first entry)  
 XX Human derived thrombomodulin.  
 DE  
 XX Human thrombomodulin; hepatopathy; fulminant hepatitis.  
 KW Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FT Protein 349..463  
 FT /note= "fragment AAR84184"  
 XX  
 XX MO9528953-A1.  
 XX  
 XX 02-NOV-1995.  
 XX  
 XX 10-APR-1995; 95WO-JP000704.  
 XX  
 XX 20-APR-1994; 94JP-00081196.  
 XX  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 XX  
 XX Fujiwara K, Mochida S;  
 XX  
 XX WPI; 1995-302844/49.  
 XX  
 XX Hepatopathy treatment pref. comprising human derived thrombomodulin -  
 FT useful against, e.g, fulminant hepatitis.  
 XX  
 XX Claim 2; Page 17-20; 29pp; Japanese.  
 XX  
 XX The human derived thrombomodulin AAR84185, and its fragment AAR84184 can  
 CC be used in the treatment of hepatopathy, including fulminant hepatitis  
 XX  
 XX Sequence 498 AA;  
 SQ

Query Match 99.9%; Score 2826; DB 2; Length 498;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-149;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYFGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 60



Db 1 APAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVILLNGD 60  
 QY 61 GGVGRRRLWIGLQLPFGCGDPKRLGFGFOWTGDNNTSYSRWARLDNGAPLCPLCV 120  
 Db 61 GGVGRRRLWIGLQLPFGCGDPKRLGFGFOWTGDNNTSYSRWARLDNGAPLCPLCV 120  
 QY 121 AVSAAEATVPSEPIWEQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 121 AVSAAEATVPSEPIWEQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGGEHACNA 240  
 Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGGEHACNA 240  
 QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVFNPDQPGSYSCMCETGYRLAA 300  
 Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVFNPDQPGSYSCMCETGYRLAA 300  
 QY 301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGCEVPEVDPFCFRANCEYQC 360  
 Db 301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGCEVPEVDPFCFRANCEYQC 360  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECPGYILLDDGFI 420  
 Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECPGYILLDDGFI 420  
 QY 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALARRHIGTDCDCKVGDGSGSGEPPPS 480  
 Db 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALARRHIGTDCDCKVGDGSGSGEPPPS 480  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 481 PTPGSTLTTPPAVGLVHSG 498

## RESULT 11

AA01600  
 ID AA01600 standard; protein; 498 AA.  
 XX AA01600;  
 XX 17-MAR-1997 (first entry)  
 DT Thrombomodulin TWE456 protein.  
 DE Thrombomodulin; nervous interference; treatment; medulla injury.  
 XX Homo sapiens.  
 OS JP08283174-A.  
 XX 29-OCT-1996.  
 XX 11-APR-1995; 95JP-00085202.  
 XX 11-APR-1995; 95JP-00085202.  
 XX (ASAH) ASAH KASEI KOGYO KK.  
 XX WPI; 1997-017314/02.  
 XX Treating agent for nervous interference after medulla injury - contains  
 PT thrombomodulin as active component.  
 XX Claim 2; Page 5-6; 6pp; Japanese.

CC AA01599-600 are fragments of human thrombomodulin. They can be used in  
 CC treatment of nervous interference after medulla injury. In an example a  
 CC model of medulla injury was prep. by compressing a male Wistar rat of  
 CC 200-250 g body wt., with a 20g weight for 20 mins. TMD123 dissolved in  
 CC physiological saline water was injected into the tail vein 30 mins.  
 CC before injury. Movement function was evaluated 24 hrs. after injury. MPO

CC activity in medulla tissue was measured 3 hrs. after injury as an index  
 CC of neutrophile accumulation. The Tarlov score was 1.50 compared to 0.77  
 CC for the control, and complete paralysis was 15.4 percent compared to 35.3  
 CC percent for the control  
 XX Sequence 498 AA;

Query Match 99.9%; Score 2826; DB 2; Length 498;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-149;  
 Matches 497; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 APAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVILLNGD 60  
 Db 1 APAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGLMTVRSSVAADVILLNGD 60  
 QY 61 GGVGRRRLWIGLQLPFGCGDPKRLGFGFOWTGDNNTSYSRWARLDNGAPLCPLCV 120  
 Db 61 GGVGRRRLWIGLQLPFGCGDPKRLGFGFOWTGDNNTSYSRWARLDNGAPLCPLCV 120  
 QY 121 AVSAAEATVPSEPIWEQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 121 AVSAAEATVPSEPIWEQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTFFA 180  
 QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGGEHACNA 240  
 Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGGEHACNA 240  
 QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVFNPDQPGSYSCMCETGYRLAA 300  
 Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVFNPDQPGSYSCMCETGYRLAA 300  
 QY 301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGCEVPEVDPFCFRANCEYQC 360  
 Db 301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNYDLVDGCEVPEVDPFCFRANCEYQC 360  
 QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECPGYILLDDGFI 420  
 Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECPGYILLDDGFI 420  
 QY 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALARRHIGTDCDCKVGDGSGSGEPPPS 480  
 Db 421 CTDIDECENGFCGVCCHNLPGTFECICGPDSSALARRHIGTDCDCKVGDGSGSGEPPPS 480  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 481 PTPGSTLTTPPAVGLVHSG 498

## RESULT 12

AA67402  
 ID AA67402 standard; protein; 498 AA.  
 XX AA67402;  
 XX 12-MAY-2000 (first entry)  
 DT Novel sugar chain-bonded thrombomodulin-like peptide #2.  
 DE Thrombomodulin; sugar chain; thrombin; coagulation; protein C; DiC;  
 XX anticoagulant; blood coagulation disorder; heart muscle; brain; liver;  
 KW disseminated intravenous coagulation; peripheral blood vessel sclerosis;  
 KW kidney infarction; toxemia; pregnancy.  
 XX Homo sapiens.  
 OS Key Location/Qualifiers  
 FH Key Misc-difference 455  
 FT /label= Val or Ala  
 XX WC200000516-A1.  
 XX 06-JAN-2000.  
 XX

PF 30-JUN-1999; 99WC-JP003532.  
 XX 30-JUN-1998; 98JP-00199654.  
 PR (MOCH ) MOCHIDA PHARM CO LTD.  
 XX Hase S, Wakabayashi H;  
 XX WPI; 2000-160668/14.  
 XX Sugar chain-bound thrombomodulin-like peptide, for treatment of, e.g.  
 PT disseminated intravenous coagulation.  
 PT  
 XX Claim 2; Page 35-37; 40pp; Japanese.  
 XX The invention provides a thrombomodulin-like peptide bearing a glucuronic  
 CC acid(betal-3)galactose(betal-3)galactose(betal-4)xylose sugar chain  
 CC sulfated at the 3-position. The peptide binds to anti-human natural  
 CC killer cell 1 antibody. It also binds to thrombin, inhibiting coagulation  
 CC and promoting protein C activation. The peptide may be used as an  
 CC anticoagulant for the treatment of blood coagulation disorders.  
 CC disseminated intravenous coagulation (DIC), peripheral blood vessel  
 CC sclerosis, heart muscle, brain, liver or kidney infarction, and toxemia  
 CC of pregnancy. The peptide has low toxicity. The present sequence  
 CC represents a novel thrombomodulin-like peptide sequence  
 XX  
 XX Sequence 498 AA;  
 SQ  
 Query Match 99.9%; Score 2826; DB 3; Length 498;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-149;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 APAPPPGSGQCVHPCFALYPGPATFNLASQICDGLRGHLMTVRSVAADVLSILLNGD 60  
 DB 1 APAPPPGSGQCVHPCFALYPGPATFNLASQICDGLRGHLMTVRSVAADVLSILLNGD 60  
 QY 61 GGVRRRLWIGLQLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
 DB 61 GGVRRRLWIGLQLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
 QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASVITGTPFA 180  
 DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASVITGTPFA 180  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 QY 241 IPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRLAA 300  
 DB 241 IPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRLAA 300  
 QY 301 DQHRCEVDVDCILEPSPC 360  
 DB 301 DQHRCEVDVDCILEPSPC 360  
 QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNOTACPADCDPNTQASCECPGYLLDDGFI 420  
 DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNOTACPADCDPNTQASCECPGYLLDDGFI 420  
 QY 421 CTDIDECEGFGCSGVCHNLPGTFECICGPDGALAHRTGTCDSKVDGSGSGSGGPPPS 480  
 DB 421 CTDIDECEGFGCSGVCHNLPGTFECICGPDGALAHRTGTCDSKVDGSGSGSGGPPPS 480  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 481 PTPGSTLTTPPAVGLVHSG 498  
 RESULT 13  
 AAR22016  
 ID AAR22016 standard; protein; 516 AA.  
 XX

AC AAR22016;  
 XX 25-MAR-2003 (revised)  
 DT 03-JUL-1992 (first entry)  
 XX Truncated human thrombomodulin encoded by plasmid pSV2TMD1.  
 DE  
 XX Thrombin binding site; blood clotting; TMD1 deleter.  
 XX Homo sapiens.  
 OS  
 XX EP474273-A.  
 PN  
 XX 11-MAR-1992.  
 PD  
 XX 05-AUG-1991; 91EP-00202009.  
 PF  
 XX 03-AUG-1990; 90JP-00204978.  
 PR  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 XX Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;  
 PI  
 XX WPI; 1992-081820/11.  
 DR  
 XX New polypeptide inhibits blood coagulation and platelet aggregation -  
 PT promotes thrombin catalysed protein C activation for treating myocardial  
 PT infarction, thrombosis, embolism, etc.  
 PT  
 XX Example 1; Page 17; 112pp; English.  
 PS  
 XX Plasmid M13mp19TMD3 (see AAR22013) was used as template for site-directed  
 CC mutagenesis. A 177bp fragment was deleted using the "TMD1 deleter" -  
 CC oligonucleotide to give plasmid M13TMD1 which encodes the first 516 N-  
 CC terminal amino acids of human thrombomodulin. Plasmid M13TMD1 was  
 CC completely digested with HindIII and BamHI and a TMD1 fragment of ca.  
 CC 1700bp was isolated. The fragment was ligated to HindIII- and BglII-cut  
 CC plasmid pSV2-dhfr to give the recombinant plasmid pSV2TMD1. See AAR22014-  
 CC R22022 and AAR22072. (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 SQ Sequence 516 AA;  
 Query Match 99.9%; Score 2826; DB 2; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-149;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 APAPPPGSGQCVHPCFALYPGPATFNLASQICDGLRGHLMTVRSVAADVLSILLNGD 60  
 DB 19 APAPPPGSGQCVHPCFALYPGPATFNLASQICDGLRGHLMTVRSVAADVLSILLNGD 78  
 QY 61 GGVRRRLWIGLQLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
 DB 79 GGVRRRLWIGLQLPPGCGDPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASVITGTPFA 180  
 DB 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASVITGTPFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGCEHACNA 258  
 QY 241 IPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRLAA 300  
 DB 259 IPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRLAA 318  
 QY 301 DQHRCEVDVDCILEPSPC 360  
 DB 319 DQHRCEVDVDCILEPSPC 378  
 QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNOTACPADCDPNTQASCECPGYLLDDGFI 420  
 DB 379 QPLNQTSYLCVCAEGFAPIPHEPHRCMFCNOTACPADCDPNTQASCECPGYLLDDGFI 438

QY 421 CTDIDECENGFCGVCNHLPGTFECICGPDGSAARHIGTDCSGKVDGSGSGEPPPS 480  
 DB 439 CTDIDECENGFCGVCNHLPGTFECICGPDGSAARHIGTDCSGKVDGSGSGEPPPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 14  
 AAY09347  
 ID AAY09347 standard; protein; 516 AA.  
 AC AAY09347;  
 XX AAY09347;  
 DT 08-JUL-1999 (first entry)  
 DE Human thrombomodulin SEQ ID NO:1.  
 XX Human; thrombomodulin; aqueous parenteral solution; storage;  
 KW distribution; acute coronary syndrome; thrombosis; embolism; diabetes.  
 XX Homo sapiens.  
 XX WO9918994-A1.  
 PN 22-APR-1999.  
 PD 13-OCT-1998; 98WO-JP004609.  
 PF 15-OCT-1997; 97JP-00281659.  
 PR 11-NOV-1997; 97JP-00308523.  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 PA Yui M, Yokozawa A, Murata T, Tsuruta K, Shimizu H;  
 PI WPI; 1999-277444/23.  
 DR N-PSDB; AAX55879.  
 XX Stable aqueous parenteral thrombomodulin solution - comprising buffer and  
 PT surfactant, useful for treating acute coronary syndrome, thrombosis,  
 PT embolism, and diabetes.  
 PS Claim 6; Page 87-89; 97pp; Japanese.  
 XX The present invention describes a method for maintaining the quality of  
 CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
 CC surfactant aseptically filled in a case or syringe. Maintaining the  
 CC quality of an aqueous, parenteral thrombomodulin solution is  
 CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
 CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
 CC filled into: (i) a case or (ii) a syringe without any empty space; and  
 CC (c) is kept in liquid form in storage and distribution and not frozen or  
 CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
 CC (such as myocardial infarction, unstable angina and coronary artery  
 CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
 CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
 CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic  
 CC lupus erythematosus or Barrett's syndrome), multiple organ failure,  
 CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
 CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
 CC is stable for a long period of time and can be stored and distributed in  
 CC ready to use form avoiding the problems of dissolution and accuracy when  
 CC preparing on demand. The present sequence represents human thrombomodulin  
 XX Sequence 516 AA;

Query Match 99.9%; Score 2826; DB 2; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-149;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCVHEHDFALYPGPATFLNASQICDGLRHLMTVRSSVAADVISLLNGD 60  
 DB 19 APAEPQGGSCVHEHDFALYPGPATFLNASQICDGLRHLMTVRSSVAADVISLLNGD 78  
 QY 61 GGVGRRRLWIGLQPPCGDPPKRLGRLRGFWTGTGNNNTSYSRWRLDINGAPLCGLCV 120  
 DB 79 GGVGRRRLWIGLQPPCGDPPKRLGRLRGFWTGTGNNNTSYSRWRLDINGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEHQCEVYKADGFLCEFHPPATCRPLAVEFGAAAVSITYGTPTA 180  
 DB 139 AVSAAEATVPSEPIWEHQCEVYKADGFLCEFHPPATCRPLAVEFGAAAVSITYGTPTA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHAREAPGAWDCSVENGCGEACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHAREAPGAWDCSVENGCGEACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
 DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318  
 QY 301 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNVLDVGECEVPPDPFCFRANCEYQC 360  
 DB 319 DQHRCEVDVDCILEPSPCPORCVNTQGGFECHCYPNVLDVGECEVPPDPFCFRANCEYQC 378  
 QY 361 QPLNQTSLVCAEGFAPIPHEHRCQMFNCQTACADCPNTOASCECEGYIIDDGEI 420  
 DB 379 QPLNQTSLVCAEGFAPIPHEHRCQMFNCQTACADCPNTOASCECEGYIIDDGEI 438  
 QY 421 CTDIDECENGFCGVCNHLPGTFECICGPDGSAARHIGTDCSGKVDGSGSGEPPPS 480  
 DB 439 CTDIDECENGFCGVCNHLPGTFECICGPDGSAARHIGTDCSGKVDGSGSGEPPPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 DB 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 15  
 AAY83934  
 ID AAY83934 standard; protein; 516 AA.  
 AC AAY83934;  
 XX AAY83934;  
 DT 28-JUL-2000 (first entry)  
 DE Human thrombomodulin TMD protein.  
 XX Human; thrombomodulin; vasculitis; protein C; thrombin.  
 OS Homo sapiens.  
 XX JP2000053582-A.  
 XX 22-FEB-2000.  
 XX 06-AUG-1998; 98JP-00222688.  
 XX 06-AUG-1998; 98JP-00222688.  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 XX WPI; 2000-353249/31.  
 XX N-PSDB; AAA10027.  
 XX Treating agent for vasculitis contains peptide which promotes activation  
 XX of protein C by thrombin.  
 XX Claim 4; Page 10-12; 18pp; Japanese.  
 XX This sequence represents a human thrombomodulin protein. The invention  
 XX relates to a treating agent for vasculitis containing a peptide which  
 XX promotes activation of protein C by thrombin

|  |     |  |
|--|-----|--|
| SQ Sequence 516 AA;  |     |  |
| Query Match 99.9%; Score 2826; DB 3; Length 516;             |     |  |
| Best Local Similarity 99.8%; Pred. No. 1.3e-149;             |     |  |
| Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0; |     |  |
| Qy   | 1   | APAEPOGSGQVEHDCFALYGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60      |
| Db   | 19  | APAEPOGSGQVEHDCFALYGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78      |
| Qy   | 61  | GGVGRRLWIGLQPLPGCGDKRLGLRGPQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120     |
| Db   | 79  | GGVGRRLWIGLQPLPGCGDKRLGLRGPQWVTGDNNTSYSRWARLDLNGAPLCPLCV 138     |
| Qy   | 121 | AVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCEPLAVEPGAAAAVSITYGTFFA 180     |
| Db   | 139 | AVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCEPLAVEPGAAAAVSITYGTFFA 198     |
| Qy   | 181 | ARGADFOALFVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSGEHACNA 240 |
| Db   | 199 | ARGADFOALFVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSGEHACNA 258 |
| Qy   | 241 | IPGAPRCQCPAGALQADGRSCTASATOSCNDLCEHFCVNPDPQPGSYSCWCETGYRLAA 300  |
| Db   | 259 | IPGAPRCQCPAGALQADGRSCTASATOSCNDLCEHFCVNPDPQPGSYSCWCETGYRLAA 318  |
| Qy   | 301 | DQHRCEVDVDCILESPQPCQRCVNTQGGFECHYCPNVDLVDGECVEPVDFCFRANCEYQC 360 |
| Db   | 319 | DQHRCEVDVDCILESPQPCQRCVNTQGGFECHYCPNVDLVDGECVEPVDFCFRANCEYQC 378 |
| Qy   | 361 | QPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACQACPADCDPNTQASCECEGYLDDGFI 420   |
| Db   | 379 | QPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACQACPADCDPNTQASCECEGYLDDGFI 438   |
| Qy   | 421 | CTDIDECENGCGFCGVCHNLPGTFECICGPDALAHIGTCDSDGKVDGDSGSGGEPPPS 480   |
| Db   | 439 | CTDIDECENGCGFCGVCHNLPGTFECICGPDALAHIGTCDSDGKVDGDSGSGGEPPPS 498   |
| Qy   | 481 | PTPGSTLTTPPAVGLVHSG 498  |
| Db   | 499 | PTPGSTLTTPPAVGLVHSG 516  |

Search completed: June 9, 2004, 08:47:58  
Job time : 54.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:46:09 ; Search time 16.5 Seconds  
(without alignments)  
1558.166 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516

Perfect score: 2830  
Sequence: 1 APABPQPGSQVCHDFAL.....PGPTPGSTLTPPAVLGHSG 498

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCRTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description                           |
|------------|--------|-------------|--------|-------|---------------------------------------|
| 1          | 2830   | 100.0       | 575    | 1     | US-08-312-870-1                       |
| 2          | 2830   | 100.0       | 575    | 6     | Sequence 1, Appli Patent No. 5466668  |
| 3          | 2826   | 99.9        | 498    | 2     | Sequence 2, Appli                     |
| 4          | 2826   | 99.9        | 575    | 1     | US-08-733-564-2                       |
| 5          | 2826   | 99.9        | 575    | 4     | Sequence 59, Appli                    |
| 6          | 2824   | 98.8        | 497    | 1     | US-08-261-206A-59                     |
| 7          | 2822   | 99.7        | 575    | 1     | US-09-880-484D-2                      |
| 8          | 2820   | 99.6        | 497    | 4     | Sequence 3, Appli                     |
| 9          | 2760.5 | 97.5        | 572    | 6     | Sequence 54, Appli Patent No. 5256770 |
| 10         | 2680   | 94.7        | 476    | 1     | US-08-014-723-1                       |
| 11         | 2680   | 94.7        | 476    | 1     | US-08-110-011A-1                      |
| 12         | 2680   | 94.7        | 494    | 1     | Sequence 1, Appli                     |
| 13         | 2680   | 94.7        | 494    | 1     | US-08-014-723-14                      |
| 14         | 2678   | 94.6        | 476    | 1     | Sequence 14, Appli                    |
| 15         | 2678   | 94.6        | 476    | 1     | Sequence 14, Appli                    |
| 16         | 2678   | 94.6        | 476    | 1     | Sequence 2, Appli                     |
| 17         | 2678   | 94.6        | 476    | 1     | US-08-014-723-18                      |
| 18         | 2678   | 94.6        | 476    | 1     | US-08-110-011A-2                      |
| 19         | 2678   | 94.6        | 494    | 1     | Sequence 18, Appli                    |
| 20         | 2678   | 94.6        | 494    | 1     | US-08-110-011A-18                     |
| 21         | 2604   | 92.0        | 475    | 1     | Sequence 16, Appli                    |
| 22         | 2604   | 92.0        | 475    | 1     | US-08-014-723-16                      |
| 23         | 2600   | 91.9        | 475    | 1     | Sequence 16, Appli                    |
| 24         | 2596   | 91.7        | 456    | 1     | US-08-110-011A-16                     |
| 25         | 2596   | 91.7        | 456    | 1     | US-08-307-444A-2                      |
| 26         | 2592   | 91.6        | 456    | 1     | Sequence 2, Appli                     |
| 27         | 2592   | 91.6        | 456    | 1     | US-08-587-389-2                       |
|            |        |             |        |       | Sequence 2, Appli                     |
|            |        |             |        |       | US-08-307-444A-1                      |
|            |        |             |        |       | US-08-587-389-1                       |
|            |        |             |        |       | US-08-307-444A-4                      |
|            |        |             |        |       | US-08-587-389-4                       |
|            |        |             |        |       | US-08-307-444A-3                      |
|            |        |             |        |       | US-08-587-389-3                       |

|    |       |      |      |   |                   |
|----|-------|------|------|---|-------------------|
| 28 | 2543  | 89.9 | 446  | 1 | US-08-307-444A-5  |
| 29 | 2543  | 89.9 | 446  | 1 | US-08-587-389-5   |
| 30 | 1621  | 57.3 | 275  | 1 | US-08-312-870-7   |
| 31 | 1159  | 41.0 | 215  | 1 | US-08-312-870-5   |
| 32 | 689   | 24.3 | 115  | 1 | US-08-312-870-9   |
| 33 | 681   | 24.1 | 114  | 2 | US-08-733-564-1   |
| 34 | 579   | 20.5 | 652  | 2 | US-08-751-305-2   |
| 35 | 566.5 | 20.0 | 492  | 4 | US-09-724-864-39  |
| 36 | 356   | 12.6 | 638  | 2 | US-08-897-443-1   |
| 37 | 352   | 12.4 | 58   | 1 | US-08-261-206A-3  |
| 38 | 348.5 | 12.3 | 915  | 4 | US-09-507-794A-34 |
| 39 | 348.5 | 12.3 | 915  | 4 | US-09-905-125A-34 |
| 40 | 348.5 | 12.3 | 915  | 4 | US-09-902-775A-34 |
| 41 | 334.5 | 11.8 | 1964 | 4 | US-09-467-997-1   |
| 42 | 326.5 | 11.5 | 956  | 2 | US-08-897-443-3   |
| 43 | 323   | 11.4 | 1253 | 3 | US-08-897-443-3   |
| 44 | 322.5 | 11.4 | 1394 | 6 | US-08-479-722B-4  |
| 45 | 321   | 11.3 | 2321 | 4 | US-09-230-652-2   |

## ALIGNMENTS

### RESULT 1

US-08-312-870-1  
; Sequence 1, Application US/08312870  
; Patent No. 5639625  
; GENERAL INFORMATION:  
; APPLICANT: Carson, Craig W.  
; APPLICANT: Esmon, Charles T.  
; TITLE OF INVENTION: Method for Detecting Antibodies to  
; TITLE OF INVENTION: Thrombomodulin in Patients  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Richards, Medlock & Andrews  
; STREET: 1201 Elm Street, Suite 4500  
; CITY: Dallas  
; STATE: Texas  
; COUNTRY: US  
; ZIP: 75270-2197  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/312,870  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hansen, Eugenia S.  
; REGISTRATION NUMBER: 31,966  
; REFERENCE/DOCKET NUMBER: OMRF B35150  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 214-939-4500  
; TELEFAX: 214-939-4600  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 19..575  
; US-08-312-870-1

Query Match 100.0%; Score 2830; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 7.5e-193;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 60  
DB 19 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 78  
QY 61 GGVRRRLWIGLQPLPGCGDKRILGRLGFWNTGDNNTSYSRWRLDNLGAPLCPLCV 120  
DB 79 GGVRRRLWIGLQPLPGCGDKRILGRLGFWNTGDNNTSYSRWRLDNLGAPLCPLCV 138  
QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEFGAAAAVSIYTGTPFA 180  
DB 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEFGAAAAVSIYTGTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318  
QY 301 DOHRCEDVDCLLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYOC 360  
DB 319 DOHRCEDVDCLLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYOC 378  
QY 361 OPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
DB 379 OPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
QY 421 CTDIDCEGNGFCVGVCHNLPGTFECICGPDLSALARIHGTDCDCKVDGSGSGSEPPPS 480  
DB 439 CTDIDCEGNGFCVGVCHNLPGTFECICGPDLSALARIHGTDCDCKVDGSGSGSEPPPS 498  
QY 481 PTFGSTLTTPPAVGLVHSG 498  
DB 499 PTFGSTLTTPPAVGLVHSG 516  
RESULT 2  
546668-6  
; Patent No. 546668  
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,  
; DAVID R.  
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR  
; PHARMACEUTICAL USE  
; NUMBER OF SEQUENCES: 57  
; CURRENT APPLICATION DATA:  
; FILING DATE: 22-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 568,456  
; FILING DATE: 15-AUG-1990  
; APPLICATION NUMBER: 506,325  
; FILING DATE: 09-APR-1990  
; APPLICATION NUMBER: 406,941  
; FILING DATE: 13-SEP-1989  
; APPLICATION NUMBER: 345,374  
; FILING DATE: 28-APR-1989  
; SEQ ID NO: 6:  
; LENGTH: 575  
546668-6

Query Match 100.0%; Score 2830; DB 6; Length 575;  
Best Local Similarity 100.0%; Pred. No. 7.5e-193;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 60  
DB 19 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSRSSVAADVISLLNGD 78  
QY 61 GGVRRRLWIGLQPLPGCGDKRILGRLGFWNTGDNNTSYSRWRLDNLGAPLCPLCV 120  
DB 79 GGVRRRLWIGLQPLPGCGDKRILGRLGFWNTGDNNTSYSRWRLDNLGAPLCPLCV 138

QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEFGAAAAVSIYTGTPFA 180  
DB 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEFGAAAAVSIYTGTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
DB 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 258  
QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318  
QY 301 DOHRCEDVDCLLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYOC 360  
DB 319 DOHRCEDVDCLLEPSPQPCQRCVNTQGGFECHCYPNYDLVDGECVPEVDPCEFRANCEYOC 378  
QY 361 OPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 420  
DB 379 OPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILDDGFI 438  
QY 421 CTDIDCEGNGFCVGVCHNLPGTFECICGPDLSALARIHGTDCDCKVDGSGSGSEPPPS 480  
DB 439 CTDIDCEGNGFCVGVCHNLPGTFECICGPDLSALARIHGTDCDCKVDGSGSGSEPPPS 498  
QY 481 PTFGSTLTTPPAVGLVHSG 498  
DB 499 PTFGSTLTTPPAVGLVHSG 516  
RESULT 3  
US-08-733-564-2  
; Sequence 2, Application US/08733564  
; Patent No. 5916874  
; GENERAL INFORMATION:  
; APPLICANT: FUJIMURA, Kenji  
; APPLICANT: MOCHIDA, Satoshi  
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
; STREET: P.O. Box 747  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/733,564  
; FILING DATE: 18 OCTOBER 1996  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SVENSSON, Leonard R.  
; REGISTRATION NUMBER: 30,330  
; REFERENCE/DOCKET NUMBER: 0216-0362P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 205-8000  
; TELEFAX: (703) 205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 498 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Protein  
US-08-733-564-2  
Query Match 99.9%; Score 2826; DB 2; Length 498;  
Best Local Similarity 99.8%; Pred. No. 1.2e-192;

Wed Jun 9 13:16:02 2004

Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 APAEPQPGSGQVCHDCFALYPGATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60

1 APAEPQPGSGQVCHDCFALYPGATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60

61 GGVGRRRLWIGLQPPGCGDKRLGLRGLFQVWTDNNTSYSRWARLDLNGAPLCPV 120

61 GGVGRRRLWIGLQPPGCGDKRLGLRGLFQVWTDNNTSYSRWARLDLNGAPLCPV 120

121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHPFATCPRLAVEPGAAAASVITYGTTPA 180

121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHPFATCPRLAVEPGAAAASVITYGTTPA 180

181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHACNA 240

181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHACNA 240

241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300

241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300

301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVPEVDFCFRANCEYQC 360

301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVPEVDFCFRANCEYQC 360

361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYILDDGFI 420

361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYILDDGFI 420

421 CTDIDECENGFCGVCVCHNLPCTFECICGPDSSALAHIGTDCDVGKVDGSGSGSGEPSPS 480

421 CTDIDECENGFCGVCVCHNLPCTFECICGPDSSALAHIGTDCDVGKVDGSGSGSGEPSPS 480

481 PTPGSTLTTPPAVGLVHSG 498

481 PTPGSTLTTPPAVGLVHSG 498

RESULT 4

US-08-261-206A-59

Sequence 59, Application US/08261206A

Patent No. 5574007

GENERAL INFORMATION:

APPLICANT: Zushi, Mitichitaka

APPLICANT: Gomi, Komakazu

APPLICANT: Yamamoto, Shuji

APPLICANT: Suzuki, Koji

APPLICANT: Matsuda, Akio

TITLE OF INVENTION: A polypeptide Capable of Interacting

TITLE OF INVENTION: with Thrombin

NUMBER OF SEQUENCES: 80

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 301 N. Washington St.

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22046-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/261,206A

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/740,492

FILING DATE: 03-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Svensson, Leonard R.

REGISTRATION NUMBER: 30330

REFERENCE/DOCKET NUMBER: 216-275P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:

LENGTH: 575 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Protein

LOCATION: 1-575

OTHER INFORMATION: /label= protein

OTHER INFORMATION: /note= "human thrombomodulin"

US-08-261-206A-59

Query Match 99.9%; Score 2826; DB 1; Length 575;

Best Local Similarity 99.8%; Pred. No. 1.4e-192;

Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 APAEPQPGSGQVCHDCFALYPGATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60

19 APAEPQPGSGQVCHDCFALYPGATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78

61 GGVGRRRLWIGLQPPGCGDKRLGLRGLFQVWTDNNTSYSRWARLDLNGAPLCPV 120

79 GGVGRRRLWIGLQPPGCGDKRLGLRGLFQVWTDNNTSYSRWARLDLNGAPLCPV 138

121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHPFATCPRLAVEPGAAAASVITYGTTPA 180

139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHPFATCPRLAVEPGAAAASVITYGTTPA 198

181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHACNA 240

199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGEHACNA 258

241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300

259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318

301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVPEVDFCFRANCEYQC 360

319 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVPEVDFCFRANCEYQC 378

361 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYILDDGFI 420

379 QPLNQTSYLVCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCEPGEYILDDGFI 438

421 CTDIDECENGFCGVCVCHNLPCTFECICGPDSSALAHIGTDCDVGKVDGSGSGSGEPSPS 480

439 CTDIDECENGFCGVCVCHNLPCTFECICGPDSSALAHIGTDCDVGKVDGSGSGSGEPSPS 498

481 PTPGSTLTTPPAVGLVHSG 498

499 PTPGSTLTTPPAVGLVHSG 516

RESULT 5

US-09-880-484D-2

Sequence 2, Application US/09880484D

Patent No. 6632791

GENERAL INFORMATION:

APPLICANT: Light, David

APPLICANT: Nagashima, Mariko

APPLICANT: Morser, Michael J

TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use

FILE REFERENCE: 51863AUSM1

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/312,870  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hansen, Eugenia S.  
REGISTRATION NUMBER: 31,966  
REFERENCE/DOCKET NUMBER: OMRF B35150  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 214-939-4500  
TELEFAX: 214-939-4600  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-312-870-3

Query Match 99.9%; Score 2826; DB 4; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.4e-192;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 19 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 78  
QY 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
Db 79 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 258  
QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
Db 259 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 378  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
Db 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 438  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

Query Match 99.8%; Score 2824; DB 1; Length 497;  
Best Local Similarity 100.0%; Pred. No. 1.7e-192;  
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
QY 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
Db 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
Db 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
Db 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
Db 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
QY 481 PTPGSTLTTPPAVGLVHSG 497  
Db 481 PTPGSTLTTPPAVGLVHSG 497

Query Match 99.9%; Score 2826; DB 4; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.4e-192;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 19 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 78  
QY 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
Db 79 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 258  
QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
Db 259 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 378  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
Db 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 438  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

Query Match 99.8%; Score 2824; DB 1; Length 497;  
Best Local Similarity 100.0%; Pred. No. 1.7e-192;  
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
QY 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
Db 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
Db 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
Db 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
Db 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
Db 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
QY 481 PTPGSTLTTPPAVGLVHSG 497  
Db 481 PTPGSTLTTPPAVGLVHSG 497

Query Match 99.9%; Score 2826; DB 4; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.4e-192;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 19 APAEPQGGSCVHDCFCALYFGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 78  
QY 61 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 120  
Db 79 GGVRRRLMIGLQLPKCGDPEKLGRLGFQWVTGDNNTSYSEWARLDLNGAPLCGLCV 138  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAVEPGAAAAVSYTGTPFA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQIMCTAPGAVQGHWAREAPGAWDCSVNGGCEHACNA 258  
QY 241 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 300  
Db 259 IPGAPROCPCAGALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYCMCTGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQPCQVNTQGGFCHCPVNDLVGECVEPVDFCFRANCEYQC 378  
QY 361 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 420  
Db 379 QPLNQTSLVCVCAEGFAPIPHEPHRCQMFNCTACPADCDNTQASCCEPGYILLDDGFI 438  
QY 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 480  
Db 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALAHIGTDCDCKVGGSGSGEPSPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 7  
US-08-170-290A-54  
Sequence 54, Application US/08170290A  
Patent No. 5702931  
GENERAL INFORMATION:  
APPLICANT: Andrews, William H.  
APPLICANT: Morser, Michael J.  
APPLICANT: Zielander, Laura R.  
TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and  
TITLE OF INVENTION: Compositions  
NUMBER OF SEQUENCES: 63

RESULT 6  
US-08-312-870-3  
Sequence 3, Application US/08312870  
Patent No. 5639625  
GENERAL INFORMATION:  
APPLICANT: Carson, Craig W.  
APPLICANT: Esmon, Charles T.  
TITLE OF INVENTION: Method for Detecting Antibodies to  
TITLE OF INVENTION: Thrombomodulin in Patients  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Richards, Medlock & Andrews  
STREET: 1201 Elm Street, Suite 4500  
CITY: Dallas  
STATE: Texas  
COUNTRY: US  
ZIP: 75270-2197  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible



CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James M. Heslin  
; STREET: 379 Lytton Ave.  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/170,290A  
; FILING DATE: 28-DEC-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05573  
; FILING DATE: 01-JUL-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/724,237  
; FILING DATE: 01-JUL-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heslin, James M.  
; REGISTRATION NUMBER: 29,541  
; REFERENCE/DOCKET NUMBER: 11972-58-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-170-290A-54

Query Match 99.7%; Score 2822; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 2.8e-192;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 19 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 78  
QY 61 GGVGRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
Db 79 GGVGRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138  
QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTTPA 180  
Db 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTTPA 198  
QY 181 ARGADFOALPVGSSAAVAPLGLQMTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 240  
Db 199 ARGADFOALPVGSSAAVAPLGLQMTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 258  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318  
QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 378  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 420  
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 438  
QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSGSGSEPPPS 480

Db 439 CTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSGSGSEPPPS 498  
QY 481 PTPGSTLTTPPAVGLVHSG 498  
Db 499 PTPGSTLTTPPAVGLVHSG 516  
RESULT 8  
US-09-331-793-4  
; Sequence 4, Application US/09331793  
; Patent No. 6500646  
; GENERAL INFORMATION:  
; APPLICANT: KURIYAMA, Shinichi  
; APPLICANT: HASEGAWA, Takashi  
; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS  
; FILE REFERENCE: 1110-253P  
; CURRENT APPLICATION NUMBER: US/09/331,793  
; CURRENT FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 497  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; US-09-331-793-4

Query Match 99.6%; Score 2820; DB 4; Length 497;  
Best Local Similarity 99.8%; Pred. No. 3.3e-192;  
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGHLMTVSSVAADVLSLLNGD 60  
QY 61 GGVGRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
Db 61 GGVGRRRLWIGLQPPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120  
QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTTPA 180  
Db 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAAVSIYGTTPA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQMTAPPGAVQGHWAREAPGAWDCSVENGCEHACNA 240  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
QY 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
Db 301 DQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 420  
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYLLDDGFI 420  
QY 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSGSGSEPPPS 480  
Db 421 CTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDCDSKVDGSGSGSEPPPS 480  
QY 481 PTPGSTLTTPPAVGLVHSG 497  
Db 481 PTPGSTLTTPPAVGLVHSG 497

RESULT 9  
5256770-7  
; Patent No. 5256770  
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,  
; DAVID R.  
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS

Wed Jun 9 13:16:02 2004

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/014,723  
 FILING DATE: 19930208  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5273962man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
 TELEPHONE: (703)413-3000  
 TELEFAX: (703)413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 476 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-014-723-1

Query Match 94.7%; Score 2680; DB 1; Length 476;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-182;  
 Matches 472; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 60  
 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 60  
 61 GGVGRRRLWIGLQPLPGGDKPKRLGFGQVWTDGNNNTSYSRWARLDINGAPLCPLCV 120  
 61 GGVGRRRLWIGLQPLPGGDKPKRLGFGQVWTDGNNNTSYSRWARLDINGAPLCPLCV 120  
 121 AVSAEATVPSPPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
 121 AVSAEATVPSPPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
 181 ARGADFOALPVGSSAAVAFGLQLMCTAPPGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 181 ARGADFOALPVGSSAAVAFGLQLMCTAPPGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 241 IPGAPRCQCPAGAAQADGSCCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 241 IPGAPRCQCPAGAAQADGSCCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 301 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPDPFCFRANCEYQC 360  
 301 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPDPFCFRANCEYQC 360  
 361 QPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 420  
 361 QPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 420  
 421 CTDIDECENGFCGVCNHLPGTFECICGPDLSALRHIGTDCDSKVDGSDSGSGGE 476  
 421 CTDIDECENGFCGVCNHLPGTFECICGPDLSALRHIGTDCDSKVDGSDSGSGGE 476

RESULT 11  
 US-08-110-011A-1  
 ; Sequence 1, Application US/08110011A  
 ; Patent No. 5354664  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doi, Takeshi  
 ; APPLICANT: Iwasaki, Akio  
 ; APPLICANT: Saino, Yushi  
 ; APPLICANT: Kimura, Shigeru  
 ; APPLICANT: Ohkuchi, Masao  
 ; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
 ; TITLE OF INVENTION: For Preparing the Same  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ;

Query Match 97.5%; Score 2760.5; DB 6; Length 572;  
 Best Local Similarity 98.6%; Pred. No. 6.2e-188;  
 Matches 491; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 60  
 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRSVAADVLSLLNGD 78  
 61 GGVGRRRLWIGLQPLPGGDKPKRLGFGQVWTDGNNNTSYSRWARLDINGAPLCPLCV 120  
 79 GGVGRRRLWIGLQPLPGGDKPKRLGFGQVWTDGNNNTSYSRWARLDINGAPLCPLCV 137  
 121 AVSAEATVPSPPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 180  
 138 AVSAEATVPSPPIWEEQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTTPFA 197  
 181 ARGADFOALPVGSSAAVAFGLQLMCTAPPGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 198 ARGADFOALPVGSSAAVAFGLQLMCTA--GNVQGHWAREAPGAWDCSVENGGEHACNA 255  
 241 IPGAPRCQCPAGAAQADGSCCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 256 IPGAPRCQCPAGAAQADGSCCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 315  
 301 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPDPFCFRANCEYQC 360  
 316 DQHRCEVDVDCILPEPSPQPCQRCVNTQGGFCHCYPNYDLVDGECVEPDPFCFRANCEYQC 375  
 361 QPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 420  
 376 QPLNQTSLYLCVCAEGFAPIPHEHRCQMFNCNOTACPADCDPNTQASCECEGYILDDGFI 435  
 421 CTDIDECENGFCGVCNHLPGTFECICGPDLSALRHIGTDCDSKVDGSDSGSGGEPPPS 480  
 436 CTDIDECENGFCGVCNHLPGTFECICGPDLSALRHIGTDCDSKVDGSDSGSGGEPPPS 495  
 481 PTPGSTLTTPRAVGLVHSG 498  
 496 PTPGSTLTTPRAVGLVHSG 513

RESULT 10  
 US-08-014-723-1  
 ; Sequence 1, Application US/08014723  
 ; Patent No. 5273962  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doi, Takeshi  
 ; APPLICANT: Iwasaki, Akio  
 ; APPLICANT: Saino, Yushi  
 ; APPLICANT: Kimura, Shigeru  
 ; APPLICANT: Ohkuchi, Masao  
 ; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
 ; TITLE OF INVENTION: For Preparing the Same  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ;

Wed Jun 9 13:16:02 2004

APPLICANT: Iwasaki, Akio  
 APPLICANT: Saino, Yushi  
 APPLICANT: Kimura, Shigeru  
 APPLICANT: Ohkuchi, Masao  
 TITLE OF INVENTION: Thrombin-Binding Substance and Process  
 TITLE OF INVENTION: For Preparing the Same  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ADDRESSEE: P.C.  
 STREET: 1755 Jefferson Davis Highway, Fourth Floor  
 CITY: Arlington  
 STATE: Virginia  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/014,723  
 FILING DATE: 19930208  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5273962man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)413-3000  
 TELEFAX: (703)413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 494 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-014-723-14

Query Match 94.7%; Score 2680; DB 1; Length 494;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-182; Indels 0; Gaps 0;  
 Matches 472; Conservative 0; Mismatches 4;  
 QY 1 APAEPQPGSQCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
 DB 19 APAEPQPGSQCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 78  
 QY 61 GVGRRRLWIGLQPPGCCDPRKLGPRFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120  
 DB 79 GVGRRRLWIGLQPPGCCDPRKLGPRFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 138  
 QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 180  
 DB 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMTAPPVQGHWAREAPGAWDCSVENGCEHACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQMTAPPVQGHWAREAPGAWDCSVENGCEHACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 DB 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318  
 QY 301 DQHCEDVDCCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQC 360  
 DB 319 DQHCEDVDCCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQC 378  
 QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYILDDGFI 420  
 DB 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYILDDGFI 438  
 QY 421 CTDIDECENGFCGSGVCHNLPTFTFCICGPDSSALVRHIGTDCDSKGVDDSGSGSE 476

ADDRESSEE: P.C.  
 STREET: 1755 Jefferson Davis Highway, Fourth Floor  
 CITY: Arlington  
 STATE: Virginia  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/110,011A  
 FILING DATE: 23-AUG-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5354664man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)413-3000  
 TELEFAX: (703)413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 476 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-110-011A-1  
 Query Match 94.7%; Score 2680; DB 1; Length 476;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-182; Indels 0; Gaps 0;  
 Matches 472; Conservative 0; Mismatches 4;  
 QY 1 APAEPQPGSQCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
 DB 1 APAEPQPGSQCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
 QY 61 GVGRRRLWIGLQPPGCCDPRKLGPRFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120  
 DB 61 GVGRRRLWIGLQPPGCCDPRKLGPRFQWVTGDNNTSYSRWARLDLNGAPLCPLCV 120  
 QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 180  
 DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 180  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMTAPPVQGHWAREAPGAWDCSVENGCEHACNA 240  
 DB 181 ARGADFOALPVGSSAAVAPLGLQMTAPPVQGHWAREAPGAWDCSVENGCEHACNA 240  
 QY 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
 QY 301 DQHCEDVDCCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQC 360  
 DB 301 DQHCEDVDCCILEPSPQPCORCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQC 360  
 QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYILDDGFI 420  
 DB 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGYILDDGFI 420  
 QY 421 CTDIDECENGFCGSGVCHNLPTFTFCICGPDSSALVRHIGTDCDSKGVDDSGSGSE 476  
 DB 421 CTDIDECENGFCGSGVCHNLPTFTFCICGPDSSALVRHIGTDCDSKGVDDSGSGSE 476

RESULT 12  
 US-08-014-723-14  
 ; Sequence 14, Application US/08014723  
 ; Patent No. 5273962  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doi, Takeshi

Db 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYFNVDLVGECVDFPCFRANCEYQC 378  
Qy 361 QPLNQTSLYCVCAEGFAPIPHEPHRCQFNCQNTACPADCDPNTQASCECPGYILDDGFI 420  
Db 379 QPLNQTSLYCVCAEGFAPIPHEPHRCQFNCQNTACPADCDPNTQASCECPGYILDDGFI 438  
Qy 421 CTDIDECENGFGCSGVCHNLPFTFECICGPDALSARHIGTDCDSKVDGSGSGE 476  
Db 439 CTDIDECENGFGCSGVCHNLPFTFECICGPDALSARHIGTDCDSKVDGSGSGE 494

RESULT 14

US-08-014-723-2  
; Sequence 2, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESSEE: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-2

Query Match 94.6%; Score 2678; DB 1; Length 476;  
Best Local Similarity 98.9%; Pred. No. 3.6e-182;  
Matches 471; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
Db 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
Qy 61 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 120  
Db 61 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 120  
Qy 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 180  
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 180

439 CTDIDECENGFGCSGVCHNLPFTFECICGPDALSARHIGTDCDSKVDGSGSGE 494  
US-08-110-011A-14  
; Sequence 14, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESSEE: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-14

Query Match 94.7%; Score 2680; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.7e-182;  
Matches 472; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
Db 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 78  
Qy 61 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 120  
Db 79 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 138  
Qy 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 198  
Qy 181 ARGADFOALPVGSSAAVPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 199 ARGADFOALPVGSSAAVPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 258  
Qy 241 IPGAPRCQCPAGALQADGSCSTASQSCNDLCEHFCVNPDPQSGYSQCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGALQADGSCSTASQSCNDLCEHFCVNPDPQSGYSQCMCTGYRLAA 318  
Qy 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYFNVDLVGECVDFPCFRANCEYQC 360

Query Match 94.7%; Score 2680; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.7e-182;  
Matches 472; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 60  
Db 1 APAEPQPGSGSCVHDCFCALYPGPATFLNASQICDGLRGHLMTVSRSSVAADVLSLLNGD 78  
Qy 61 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 120  
Db 79 GGVGRRLWIGLQPLPPGCGDKRLGFLRGFWTGDNNNTSYSRWRLDNGAPLCGLCV 138  
Qy 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 180  
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSITYGTTPA 198  
Qy 181 ARGADFOALPVGSSAAVPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 240  
Db 199 ARGADFOALPVGSSAAVPLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGSCHEACNA 258  
Qy 241 IPGAPRCQCPAGALQADGSCSTASQSCNDLCEHFCVNPDPQSGYSQCMCTGYRLAA 300  
Db 259 IPGAPRCQCPAGALQADGSCSTASQSCNDLCEHFCVNPDPQSGYSQCMCTGYRLAA 318  
Qy 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHCYFNVDLVGECVDFPCFRANCEYQC 360

Query Match 94.6%; Score 2678; DB 1; Length 476;  
Best Local Similarity 99.2%; Pred. No. 3.6e-182; Indels 0; Gaps 0;  
Matches 471; Conservative

QY 1 APABPQGGSCQVEHDCFALYPGPATFELNASQICDGLRGLHMTVRSSVAADVILLNGD 60  
DB 1 APABPQGGSCQVEHDCFALYPGPATFELNASQICDGLRGLHMTVRSSVAADVILLNGD 60  
QY 61 GVGERRLWIGLQPLPPGCGDPKRLGPGFOWVTGDNNTSYRWARLDLNGAPLGLCV 120  
DB 61 GVGERRLWIGLQPLPPGCGDPKRLGPGFOWVTGDNNTSYRWARLDLNGAPLGLCV 120  
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIITYGTTPA 180  
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAVSIITYGTTPA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHACNA 240  
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHACNA 240  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
QY 301 DQHRCEVDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
DB 301 DQHRCEVDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILLDDGFI 420  
DB 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILLDDGFI 420  
QY 421 CTDIDECENGFGFCGVCHNLPGTFECICGPDLSALARIHIGTDCDSKVDGDSGSG 475  
DB 421 CTDIDECENGFGFCGVCHNLPGTFECICGPDLSALARIHIGTDCDSKVDGDSGSG 475

Search completed: June 9, 2004, 08:51:31  
Job time : 18.5 secs

QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHACNA 240  
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHACNA 240  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
DB 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300  
QY 301 DQHRCEVDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
DB 301 DQHRCEVDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360  
QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILLDDGFI 420  
DB 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPNTQASCECEGYILLDDGFI 420  
QY 421 CTDIDECENGFGFCGVCHNLPGTFECICGPDLSALARIHIGTDCDSKVDGDSGSG 476  
DB 421 CTDIDECENGFGFCGVCHNLPGTFECICGPDLSALARIHIGTDCDSKVDGDSGSG 476

RESULT 15  
US-08-014-723-18  
; Sequence 18, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19930208  
; APPLICATION NUMBER: US/08/014,723  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 469  
; OTHER INFORMATION: /note= "acidic amino acid"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 470  
; OTHER INFORMATION: /note= "acidic amino acid"  
; US-08-014-723-18

Wed Jun 9 13:16:02 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:50:04 ; Search time 42 Seconds  
(without alignments)  
3335.873 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516  
Perfect score: 2830  
Sequence: 1 APAPQGGSCVHEHDFAL.....PSPTGSLTPPAVGLVHSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues  
Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap:  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap:  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap:  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pap:  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pap:  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pap:  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pap:  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap:  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pap:  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pap:  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pap:  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pap:  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap:  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap:  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pap:  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pap:  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pap:  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pap:

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description         |
|------------|-------|-------|--------|----|---------------------|
| 1          | 2830  | 100.0 | 575    | 14 | US-10-150-440-3     |
| 2          | 2830  | 100.0 | 575    | 15 | US-10-373-801-29    |
| 3          | 2830  | 100.0 | 631    | 14 | US-10-150-440-1     |
| 4          | 2826  | 99.9  | 575    | 9  | US-09-938-405-2     |
| 5          | 2826  | 99.9  | 575    | 10 | US-09-880-464-2     |
| 6          | 2826  | 99.9  | 575    | 14 | US-10-438-648-2     |
| 7          | 2826  | 99.9  | 575    | 15 | US-10-410-195-2     |
| 8          | 2820  | 99.6  | 497    | 15 | US-10-298-796-4     |
| 9          | 2819  | 99.6  | 575    | 15 | US-10-094-886-196   |
| 10         | 1032  | 36.5  | 239    | 15 | US-10-104-047-2759  |
| 11         | 735   | 26.0  | 418    | 12 | US-10-427-805-2     |
| 12         | 715   | 25.3  | 397    | 12 | US-10-427-805-3     |
| 13         | 582   | 20.6  | 645    | 14 | US-10-029-386-33151 |
| 14         | 582   | 20.6  | 652    | 16 | US-10-408-765A-1422 |
| 15         | 580   | 20.5  | 652    | 9  | US-09-789-919-96    |

Sequence 131, App  
Sequence 83, Appl  
Sequence 176, App  
Sequence 177, App  
Sequence 196, App  
Sequence 107, App  
Sequence 190, App  
Sequence 291, App  
Sequence 62, Appl  
Sequence 73, Appl  
Sequence 77, Appl  
Sequence 60, Appl  
Sequence 482, App  
Sequence 59, Appl  
Sequence 10, Appl  
Sequence 8, Appl  
Sequence 12, Appl  
Sequence 6, Appl  
Sequence 68, Appl  
Sequence 162, App  
Sequence 421, App  
Sequence 14, Appl  
Sequence 57, Appl  
Sequence 56, Appl  
Sequence 195, App  
Sequence 21, Appl  
Sequence 102, App  
Sequence 1187, App  
Sequence 176, App  
Sequence 1292, App

ALIGNMENTS

RESULT 1  
US-10-150-440-3  
; Sequence 3, Application US/10150440  
; Publication No. US2003002249A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmitz, Juergen  
; APPLICANT: Buzek, David William  
; TITLE OF INVENTION: ANTIGEN-BINDING FRAGMENTS THAT RECOGNIZE  
; TITLE OF INVENTION: A SUBSET OF DENDRITIC CELLS AND METHODS OF USE THEREOF  
; FILE REFERENCE: 212302001200  
; CURRENT APPLICATION NUMBER: US/10/150,440  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR FILING DATE: 2002-10-01  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR FILING DATE: 1999-11-15  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-150-440-3

Query Match 100.0%; Score 2830; DB 14; Length 575;

Best Local Similarity 100.0%; Pred. No. 3.9e-193; Indels 0; Gaps 0;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
19 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 78  
61 GGVGRRRLWIGLQPLPGGDPKRLGFLRGFOWVTGDNNTSRSRWALDLNAPLCGLCV 120  
79 GGVGRRRLWIGLQPLPGGDPKRLGFLRGFOWVTGDNNTSRSRWALDLNAPLCGLCV 138  
121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 180  
139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 198  
181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVNGGCEHCACNA 240  
199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVNGGCEHCACNA 258  
241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDICEHFCVNPDPQPGSYSCMCETGYRLAA 300  
259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDICEHFCVNPDPQPGSYSCMCETGYRLAA 318  
301 DOHRCEDVDDCILEPSPQPCORCVNTQGFECCHCYPNYDLVDGECVBPVDFCFRANCEYQC 360  
319 DOHRCEDVDDCILEPSPQPCORCVNTQGFECCHCYPNYDLVDGECVBPVDFCFRANCEYQC 378  
361 QPLNQTSLVLCVACAGFAPIPHEPHRCQMFNQTACPADCDPNTQASCEPGEYILDDGFI 420  
379 QPLNQTSLVLCVACAGFAPIPHEPHRCQMFNQTACPADCDPNTQASCEPGEYILDDGFI 438  
421 CTTDIDECENGFCGVCNHLPGTFECICGPDSSALAHIGTDCDCKVGDGSGSGEPPPS 480  
439 CTTDIDECENGFCGVCNHLPGTFECICGPDSSALAHIGTDCDCKVGDGSGSGEPPPS 498  
481 PTPGSTLTTPPAVGLVHSG 498  
499 PTPGSTLTTPPAVGLVHSG 516

RESULT 3

US-10-150-440-1  
; Sequence 1, Application US/10150440  
; Publication No. US2003002249A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmitz, Juergen  
; APPLICANT: Dzionek, Andrzej  
; APPLICANT: Buck, David William  
; TITLE OF INVENTION: ANTIGEN-BINDING FRAGMENTS THAT RECOGNIZE  
; FILE REFERENCE: 212302001200  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: US 09/714,712  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: US 60/291,561  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 60/197,205  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 60/196,824  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: US 60/180,775  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: US 60/179,003  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/167,076  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/165,555  
; PRIOR FILING DATE: 1999-11-15  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 631  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-150-440-1

Query Match 100.0%; Score 2830; DB 14; Length 631;

Best Local Similarity 100.0%; Pred. No. 4.3e-193;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
75 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 134  
61 GGVGRRRLWIGLQPLPGGDPKRLGFLRGFOWVTGDNNTSRSRWALDLNAPLCGLCV 120

RESULT 2

US-10-373-801-29  
; Sequence 29, Application US/10373801  
; Publication No. US20040005644A1  
; GENERAL INFORMATION:  
; APPLICANT: Yibai Pharmaceutical (USA)  
; TITLE OF INVENTION: Method and composition for detection and treatment of breast cancer  
; FILE REFERENCE: 12399.00  
; CURRENT FILING DATE: 2003-02-27  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 29  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-373-801-29

Query Match 100.0%; Score 2830; DB 15; Length 575;

Best Local Similarity 100.0%; Pred. No. 3.9e-193;  
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 60  
19 APABPQGGSCVHDCFPALPGPATFLNASQICDGLRGLMTVRSSVAADVLSLLNGD 78  
61 GGVGRRRLWIGLQPLPGGDPKRLGFLRGFOWVTGDNNTSRSRWALDLNAPLCGLCV 120  
79 GGVGRRRLWIGLQPLPGGDPKRLGFLRGFOWVTGDNNTSRSRWALDLNAPLCGLCV 138  
121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 180  
139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAVSITYGTFFA 198



Db 259 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 318  
 QY 301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 360  
 Db 319 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 378  
 QY 361 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 420  
 Db 379 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 438  
 QY 421 CTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSKVDGSDSGSGEPPPS 480  
 Db 439 CTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSKVDGSDSGSGEPPPS 498  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 5  
 US-09-880-464-2  
 ; Sequence 2, Application US/09880464  
 ; Publication No. US20030011623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morser, Michael J  
 ; APPLICANT: Nagashima, Mariko  
 ; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use  
 ; FILE REFERENCE: Improved Thrombomodulin Analogs  
 ; CURRENT APPLICATION NUMBER: US/09/880,464  
 ; PRIOR FILING DATE: 2001-08-21  
 ; PRIOR FILING DATE: 2000-06-21  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 575  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; ORGANISM: Homo sapiens  
 US-09-880-464-2

Query Match 99.9%; Score 2826; DB 10; Length 575;  
 Best Local Similarity 99.8%; Pred. No. 7.5e-193;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQLPPGCGDPKRLGFLRGFQVWTDGNTTSYSRWARLDNGAPLCGLCV 120  
 Db 79 GGVGRRRLWIGLQLPPGCGDPKRLGFLRGFQVWTDGNTTSYSRWARLDNGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 300  
 Db 259 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 318  
 QY 301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 360  
 Db 319 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 378  
 QY 361 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 420  
 Db 379 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 438

Db 135 GGVGRRRLWIGLQLPPGCGDPKRLGFLRGFQVWTDGNTTSYSRWARLDNGAPLCGLCV 194  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 195 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 254  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 Db 255 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 314  
 QY 241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 300  
 Db 315 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 374  
 QY 301 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 360  
 Db 375 DQHRCEVDVDDCILEPSPQPCORCVNTQGGFECCHYCPNYDLVDGCEVPVDPFCFRANCEYQC 434  
 QY 361 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 420  
 Db 435 QPLNQTSTYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDENTQASCECPGYILDDGFI 494  
 QY 421 CTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSKVDGSDSGSGEPPPS 480  
 Db 495 CTDIDCEGNGFCGVCCHNLPGTFECICGPDSSALARIHGTDCDSKVDGSDSGSGEPPPS 554  
 QY 481 PTPGSTLTTPPAVGLVHSG 498  
 Db 555 PTPGSTLTTPPAVGLVHSG 572

RESULT 4  
 US-09-938-405-2  
 ; Sequence 2, Application US/09938405  
 ; Patent No. US20020111296A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pestoff, Barry W.  
 ; APPLICANT: Morser, Michael J.  
 ; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury  
 ; FILE REFERENCE: 51960AUSM1  
 ; CURRENT APPLICATION NUMBER: US/09/938,405  
 ; CURRENT FILING DATE: 2001-08-23  
 ; PRIOR APPLICATION NUMBER: 60/229,714  
 ; PRIOR FILING DATE: 2000-08-31  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 575  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; ORGANISM: Homo sapiens  
 US-09-938-405-2

Query Match 99.9%; Score 2826; DB 9; Length 575;  
 Best Local Similarity 99.8%; Pred. No. 7.5e-193;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 60  
 Db 19 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSVAADVLSLLNGD 78  
 QY 61 GGVGRRRLWIGLQLPPGCGDPKRLGFLRGFQVWTDGNTTSYSRWARLDNGAPLCGLCV 120  
 Db 79 GGVGRRRLWIGLQLPPGCGDPKRLGFLRGFQVWTDGNTTSYSRWARLDNGAPLCGLCV 138  
 QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 180  
 Db 139 AVSAAEATVPSEPIWEEOQCEVKADGFLCEHFFHPCATCRPLAVEPGAAAAVSIYGTFFA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 240  
 Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNA 258  
 QY 241 IPGAPRCQCPAGAAALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSCMCEGYRLAA 300



Wed Jun 9 13:16:02 2004

```

421 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 480
439 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 498
481 PTPGSTLTTPPAVGLVHSG 498
499 PTPGSTLTTPPAVGLVHSG 516

RESULT 6
US-10-438-648-2
; Sequence 2, Application US/10438648
; Publication No. US20030186883A1
; GENERAL INFORMATION:
; APPLICANT: Light, David
; APPLICANT: Nagaehima, Mariko
; APPLICANT: Morser, Michael J.
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use
; FILE REFERENCE: 51863AUSD1
; CURRENT APPLICATION NUMBER: US/10/438,648
; CURRENT FILING DATE: 2003-05-14
; PRIOR APPLICATION NUMBER: US 60/213,678
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 09/880,484
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-438-648-2

Query Match 99.9%; Score 2826; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 7.5e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 60
Db 19 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 78
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
QY 121 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 180
Db 139 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 258
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 300
Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 318
QY 301 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
Db 319 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 378
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 420
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 438
QY 421 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 480
Db 439 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

US-10-438-648-2
; Sequence 2, Application US/10438648
; Publication No. US20030186883A1
; GENERAL INFORMATION:
; APPLICANT: Light, David
; APPLICANT: Nagaehima, Mariko
; APPLICANT: Morser, Michael J.
; TITLE OF INVENTION: Thrombomodulin Analogs for Pharmaceutical Use
; FILE REFERENCE: 51863AUSD1
; CURRENT APPLICATION NUMBER: US/10/438,648
; CURRENT FILING DATE: 2003-05-14
; PRIOR APPLICATION NUMBER: US 60/213,678
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 09/880,484
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-438-648-2

Query Match 99.9%; Score 2826; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 7.5e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 60
Db 19 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 78
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
QY 121 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 180
Db 139 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 258
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 300
Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 318
QY 301 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
Db 319 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 378
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 420
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 438
QY 421 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 480
Db 439 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

```

```

RESULT 7
US-10-410-195-2
; Sequence 2, Application US/10410195
; Publication No. US20040002446A1
; GENERAL INFORMATION:
; APPLICANT: Festoff, Barry W.
; APPLICANT: Morser, Michael J.
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury
; FILE REFERENCE: 51960AUSM1
; CURRENT APPLICATION NUMBER: US/10/410,195
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US/09/938,405
; PRIOR FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/229,714
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-410-195-2

Query Match 99.9%; Score 2826; DB 15; Length 575;
Best Local Similarity 99.8%; Pred. No. 7.5e-193;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 60
Db 19 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISILLNGD 78
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
Db 79 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
QY 121 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 180
Db 139 AVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAVEPGAAAAVSIYGTPEA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 240
Db 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCGEHCACNA 258
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 300
Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCEYGLAA 318
QY 301 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 360
Db 319 DQHRCEVDVDCILEPSPQPCQVCNTQGGFECHCYPNYDLVDGECVPEVDPFCFRANCEYQC 378
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 420
Db 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPNTQASCECEGYILDDGFI 438
QY 421 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 480
Db 439 CTIDECENGFCGVCNHLPGTFFCICGPDALAHIGTDCDCKVGDGSGSGEPPPS 498
QY 481 PTPGSTLTTPPAVGLVHSG 498
Db 499 PTPGSTLTTPPAVGLVHSG 516

RESULT 8
US-10-298-796-4
; Sequence 4, Application US/10298796
; Publication No. US20030220490A1
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi

```

Wed Jun 9 13:16:02 2004

```

; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/10/298,796
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US/09/331,793
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-298-796-4

Query Match      99.6%; Score 2820; DB 15; Length 497;
Best Local Similarity 99.8%; Pred. No. 1.7e-192;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60
Db 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60
QY 61 GGVGRRRLWIGLQPPGCGDKPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
Db 61 GGVGRRRLWIGLQPPGCGDKPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 180
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 180
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 180
Db 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 180
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 240
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 240
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 240
Db 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 240
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
QY 301 DOHRCEDVDDCILBPSPCQPCVNTQGGFECHYCNVLDVDEGCEVPVDFPFRANCEYQC 360
Db 301 DOHRCEDVDDCILBPSPCQPCVNTQGGFECHYCNVLDVDEGCEVPVDFPFRANCEYQC 360
QY 301 DOHRCEDVDDCILBPSPCQPCVNTQGGFECHYCNVLDVDEGCEVPVDFPFRANCEYQC 360
Db 301 DOHRCEDVDDCILBPSPCQPCVNTQGGFECHYCNVLDVDEGCEVPVDFPFRANCEYQC 360
QY 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACPDACDNTQASCCEPGEYILDDGFI 420
Db 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFQACPDACDNTQASCCEPGEYILDDGFI 420
QY 421 CTDDIDECENGCGFCGCHNLPGTFECICGPDALRHIGTDCDCKGVDDGSDSGSGEPFPS 480
Db 421 CTDDIDECENGCGFCGCHNLPGTFECICGPDALRHIGTDCDCKGVDDGSDSGSGEPFPS 480
QY 481 PTPGSTLTTPPAVGLVHS 497
Db 481 PTPGSTLTTPPAVGLVHS 497

RESULT 9
US-10-094-886-196
; Sequence 196, Application US/10094886
; Publication No. US20040002120A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Vernet, Corine A.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shenoy, Suresh
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Miller, Charles
; APPLICANT: Casman, Stacie
; APPLICANT: Pena, Carol
; APPLICANT: Gangolli, Esha
; APPLICANT: Gusev, Vladimir
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Gerlach, Valerie
; APPLICANT: Pochart, Pascal
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderma, Steven
; APPLICANT: LaRoche, William
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-290 B
; CURRENT APPLICATION NUMBER: US/10/094,886
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,182
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/288,052
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/318,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/314,018
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/296,693
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/313,626
; PRIOR FILING DATE: 2001-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: Patent in 2.1
; SEQ ID NO 196
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-094-886-196

Query Match      99.6%; Score 2819; DB 15; Length 575;
Best Local Similarity 99.6%; Pred. No. 2.4e-192;
Matches 496; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 60
Db 19 APAEPQGGSCQVEHDCFALYGPATFLNASQICDGLRGLHMTVRSSVAADVVISLLNGD 78
QY 61 GGVGRRRLWIGLQPPGCGDKPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
Db 79 GGVGRRRLWIGLQPPGCGDKPKRLGFLRGFWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
QY 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 180
Db 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFFHPATCRPLAVEPFGAAAASVITYGTTPFA 198
QY 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 240
Db 199 ARGAGFOALPVGSSAAVAPLGLQMLCTAPPVQGHWAREAPGAWDCSVENGCEHCACNA 258
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300
Db 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 318
```

Wed Jun 9 13:16:02 2004

```

301 DOHRCEDVDDCILEPSPORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 360
319 DOHRCEDVDDCILEPSPORCVNTQGGFECHCYPNYDLVDGECVEPVPDPCFRANCEYQC 378
361 OPLNQTSLVCACGAFAPIPHEPFCNMFNQTACPADCDPNTQASCEGYYLDDGFI 420
379 OPLNQTSLVCACGAFAPIPHEPFCNMFNQTACPADCDPNTQASCEGYYLDDGFI 438
421 CTDIDECENGFCGSGVCHNLPGTFECICGPDALAHIGTDCDCKVDDGSGSGGEPSPS 480
439 CTDIDECENGFCGSGVCHNLPGTFECICGPDALAHIGTDCDCKVDDGSGSGGEPSPS 498
481 PTPGSTLTTPPAVLVHSG 498
499 PTPGSTLTTPPAVLVHSG 516

```

```

RESULT 10
US-10-104-047-2759
; Sequence 2759, Application US/10104047
; Publication No. US2003023632A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US2003023632A1el full length cdna
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2759
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2759

```

```

Query Match 36.5%; Score 1032; DB 15; Length 239;
Best Local Similarity 86.1%; Pred. No. 1.2e-65;
Matches 179; Conservative 0; Mismatches 1; Indels 28; Gaps 1;

QY 291 MCETGYRLAADQHRCEVDVDDCILEPSPORCVNTQGGFECHCYPNYDLVDGECVEPVPD 350
DB 1 MCETGYRLAADQHRCEVDVDDCILEPSPORCVNTQGGFECHCYPNYDLVDGECVEPVPD 60
QY 351 CFRANCEYOCQPLNQTSLVCACGAFAPIPHEPFCNMFNQTACPADCDPNTQASCECP 410
DB 61 CFRANCEYOCQPLNQTSLVCACGAFAPIPHEPFCNMFNQTACPADCDPNTQASCE-- 118
QY 411 EGYLLDDGPICTDIDECENGFCGSGVCHNLPGTFECICGPDALAHIGTDCDCKVDDG 470
DB 119 -----CHNLPGTFECICGPDALAHIGTDCDCKVDDG 152
QY 471 DSGSGEPSPSTPGSTLTTPPAVLVHSG 498
DB 153 DSGSGEPSPSTPGSTLTTPPAVLVHSG 180

```

```

RESULT 11
US-10-427-805-2
; Sequence 2, Application US/10427805
; Publication No. US20040063632A1
; GENERAL INFORMATION:
; APPLICANT: Light, David
; TITLE OF INVENTION: No. US20040063632A1el Tissue Factor Targeted Thrombomodulin Fuscio
; FILE REFERENCE: 52295AUSM1
; CURRENT APPLICATION NUMBER: US/10/427,805
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,566
; PRIOR FILING DATE: 2002-05-01

```

```

; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of scFv(TF)3e10-TM1456 fusion protein
US-10-427-805-2

Query Match 26.0%; Score 735; DB 12; Length 418;
Best Local Similarity 50.5%; Pred. No. 3e-44;
Matches 149; Conservative 18; Mismatches 55; Indels 73; Gaps 8;

QY 243 CAPRCQCPAGAAALQAD-GRSCTASATQSCNDLCEHFCVNPDPQSGYSYSCMCETGYRLAAD 301
DB 150 GAPNFMLTQPHSVGASPGKTVTISCTSSGVSASYVQVYQORPGSSPTTV-----IYED 204
QY 302 QHRCEDVDDCILEPSPORCVNT-----QGGFECHCYPNYDLV-- 340
DB 205 NHR-----PSGVDPFRFGSITDSSNSASLTISGLKTEDEADYYCQSYDNNLVVF 254
QY 341 -----DGECEVPDPCFRANCEYOCQPLNQTSLVCACG 376
DB 255 GGGTKLTVLGAAGGGGGGGGSGVPEVPDPCFRANCEYOCQPLNQTSLVCACG 314
QY 377 APIPHEPFCNMFNQTACPADCDPNTQASCEPGEYLLDDGFICTDIDECENGFCGSGV 436
DB 315 APIPHEPFCNMFNQTACPADCDPNTQASCEPGEYLLDDGFICTDIDECENGFCGSGV 374
QY 437 CHNLPGTFECICGPDALAHIGTDCDCKVDDGSGSGEPSPSTPGSTLTTPPA 491
DB 375 CHNLPGTFECICGPDALAHIGTDC-----AAAGAFVYVDP--LEPRA 417

```

```

RESULT 12
US-10-427-805-3
; Sequence 3, Application US/10427805
; Publication No. US20040063632A1
; GENERAL INFORMATION:
; APPLICANT: Light, David
; TITLE OF INVENTION: No. US20040063632A1el Tissue Factor Targeted Thrombomodulin Fuscio
; FILE REFERENCE: 52295AUSM1
; CURRENT APPLICATION NUMBER: US/10/427,805
; CURRENT FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: US 60/376,566
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of scFv(TF)3e10-TM1456delta fusio protein
US-10-427-805-3

```

```

Query Match 25.3%; Score 715; DB 12; Length 397;
Best Local Similarity 55.0%; Pred. No. 7.5e-43;
Matches 137; Conservative 14; Mismatches 38; Indels 60; Gaps 5;

QY 259 GRSCTASATQSCNDLCEHFCVNPDPQSGYSYSCMCETGYRLAADQHRCEVDVDDCILEPSPC 318
DB 164 GKTVTISCTSSGVSASYVQVYQORPGSSPTTV-----IYEDNHR-----PSGV 208
QY 319 POR---CVNT-----QGGFECHCYPNYDLV-- 340
DB 209 PDRFSGSIDTSSNSASLTISGLKTEDEADYYCQSYDNNLVWFGGGTKLVLGAAGGG 268
QY 341 -----DGECEVPDPCFRANCEYOCQPLNQTSLVCACGAFAPIPHEPFCNMFNQT 393

```

|  |     |   |                          |
|--|-----|---|--------------------------|
| QY   | 434 | SGVCHNLPGTFCICGPDSSALAHICHTDCDCKGV                                    | -----DGDGSGSGE-----P 477 |
| Db   | 440 | DSLCFNTQSGFHGCGCLPGWVLAEN-GVCTMGFVSLGPPSGPDEEDKGEKEGSTVPRAA           | 498                      |
| QY   | 478 | PPSPFTGSGTLTPPA 491   |                          |
| Db   | 499 | TASFTRGECTPKA 512   |                          |
| RESULT 14  |     |   |                          |
| US-10-408-765A-1422  |     |   |                          |
| ; Sequence 1422, Application US/10408765A                            |     |   |                          |
| ; Publication No. US20040101874A1                                    |     |   |                          |
| ; GENERAL INFORMATION:   |     |   |                          |
| ; APPLICANT: Ghosh, Soumitra S.                                      |     |   |                          |
| ; APPLICANT: Fahy, Eoin D.   |     |   |                          |
| ; APPLICANT: Zhang, Bing   |     |   |                          |
| ; APPLICANT: Gibson, Bradford W.                                     |     |   |                          |
| ; APPLICANT: Taylor, Steven W.                                       |     |   |                          |
| ; APPLICANT: Glenn, Gary M.  |     |   |                          |
| ; APPLICANT: Warnock, Dale E.  |     |   |                          |
| ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION           |     |   |                          |
| ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME       |     |   |                          |
| ; FILE REFERENCE: 660088.465   |     |   |                          |
| ; CURRENT APPLICATION NUMBER: US/10/408,765A                         |     |   |                          |
| ; CURRENT FILING DATE: 2003-04-04                                    |     |   |                          |
| ; NUMBER OF SEQ ID NOS: 3077   |     |   |                          |
| ; SOFTWARE: FastSeq for Windows Version 4.0                          |     |   |                          |
| ; SEQ ID NO 1422   |     |   |                          |
| ; LENGTH: 652  |     |   |                          |
| ; TYPE: PRT  |     |   |                          |
| ; ORGANISM: Homo sapiens   |     |   |                          |
| ; US-10-408-765A-1422  |     |   |                          |
| Query Match 20.6%; Score 582; DB 16; Length 652;                     |     |   |                          |
| Best Local Similarity 31.8%; Pred. No. 3.8e-33; Indels 126; Gaps 31; |     |   |                          |
| Matches 176; Conservative 52; Mismatches 200;                        |     |   |                          |
| QY   | 6   | ORQGSQ-----CVEHDCFALYPGPATFLNASQICDGLRHLMTVRSVAAD-----VI 53           |                          |
| Db   | 17  | OPGAGTGADTEAVVCGTACTYTAHSKLSAAEAQNCHNQGNNLAATVKSKEAQHVQRL 76          |                          |
| QY   | 54  | SLILNGDGGVGR--RLWIGLQLPPG-CGDPKELGLRGFOWVTGDNNTSYSRWARLDLN 110        |                          |
| Db   | 77  | AQLLRREALTARMSKFWIGLOREKGLDPSL--PLKGSFWVGGEETPYSNWHKELRN 134          |                          |
| QY   | 111 | GAPLCGLPLCA--VSAEAATVPE--PIWEEQQC-----EVKADGLCEPHFATCRPLA 161         |                          |
| Db   | 135 | SC--ISKRCVSLILDLSQPLLRPLKWSBGCPGSPGSGNIEGFVCKFSFKMGCRPLA 192          |                          |
| QY   | 162 | V-BPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQIMC--TAPPGAVQGH-- 217       |                          |
| Db   | 193 | LGPG-----QVITYTTFQTSSLEAVPFASAANVA-----CGEGDKDETQSHYFL 239            |                          |
| QY   | 218 | -AREAPGAWD-----CSVENGCCEHAC--NAIPGAPRCOCAPAGALQADGRS 261              |                          |
| Db   | 240 | CKEAPDVFDMGSSGPLCVSPKYGCNFFNGGCHQDCEGGGSGFLCGRCGFRLLDLVT 299          |                          |
| QY   | 262 | CTASATQSCNDLCE--HFCVNPDPQPGSYSCMCTGYRLAADQHRCEDDVDCILEPSPCP 319       |                          |
| Db   | 300 | C-ASRNPCCSFCRGATCVLGP-HGKNYTCRCQGYQLDSSQLDCVDVDEC--QDSPCA 355         |                          |
| QY   | 320 | ORCVNTQGGFECHCYPNYDLVDGECVEPDPFCFRANCEYQCOPLNQTSYLCVCAEGFAP 379       |                          |
| Db   | 356 | QECVNTPGGFRCECWGY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395           |                          |
| QY   | 380 | PHEPHRCQMFNCQTACPADCDNTQAS--CECEGYIL--DDGFICTDIDEC--ENGDFC 433        |                          |
| Db   | 396 | QECVNTPGGFRCECWGY-----CAQGC--TNTDGSFHCSEEGYVLAGEDGTCCQDVDCVGGGFLC 439 |                          |
| QY   | 434 | SGVCHNLPGTFCICGPDSSALAHICHTDCDCKGV                                    | -----DGDGSGSGE-----P 477 |

Wed Jun 9 13:16:02 2004

Job time : 44 secs

Db 440 DSLCFNTQSFHCGCLPGWILAPN-GVSCMTGMPVSLGPPSGPPDEEDKGEKSGSTVPRAA 498

QY 478 PPSPTPGSTLTTPPA 491

Db 499 TASPTRGEGTTPKA 512

RESULT 15

US-09-789-919-96

; Sequence 96, Application US/09789919

; Patent No. US20020064855A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor

; APPLICANT: Moore, Kateri

; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM

; TITLE OF INVENTION: CELLS AND USES THEREOF

; FILE REFERENCE: 2275-1-005

; CURRENT APPLICATION NUMBER: US/09/789,919

; CURRENT FILING DATE: 2001-02-21

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 96

; LENGTH: 652

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-789-919-96

Query Match 20.5%; Score 580; DB 9; Length 652;

Best Local Similarity 31.6%; Pred. No. 5.2e-33;

Matches 175; Conservative 53; Mismatches 200; Indels 126; Gaps 31;

QY 6 QPGGSG-----CUEHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAAD---VI 53

Db 17 QPGAGTCADTEAVVCGVTACTYTAHSGKLSAAEAQNHCNONGNLTATVKSKEAQQHVQVRL 76

QY 54 SILLNGDGVGR--RLWIGLQPPG-CGDPKRLGLRGLFQWVTGDNNTSYSRWALDLN 110

Db 77 AQLRREALTARMSFWIGLQREKGLDPSL--PLKGFVWGGGEDIPIYSNNWHELRN 134

QY 111 GAPLCGLCVA--VSAEATVPSE-PIWEQOC-----EVKADGFLCEHFPATCRPLA 161

Db 135 SC--ISKRCVSLLDLSQPLLPNRLPKWSEGPCSGPSGNSIEGFVKFSPKGMCRPLA 192

QY 162 V-EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQIMC-TAPPGAVQGHW-- 217

Db 193 LGGFG-----QVYTTTPTTSSLEAVPFAAANVA-----CGEGKDETQSHYFL 239

QY 218 -AREAPGAWD-----CSVENGGCEHAC-NAIPGAPRCQCPAGALQADGRS 261

Db 240 CKEKAPVDFWSSGGLCVSPKYGCFNFGGCHQDCPEGGDGSFLCGCRPGRLLDLVT 299

QY 262 CTASATOSCNDLCE--HFCVNPDPQGSYSYCMCTGYRLAADQHRCEDDVDDCILEPSPCP 319

Db 300 C-ASRNPCSSSPCRGGATCVLGP-HGKNYTCRCPOGYQLDSSQLDCVDVDEC--QDSPCA 355

QY 320 QRCVNTGGFECHCYVNYDLVDGCVBPVDFCFRANCEYOCQPLNQTSYLCVCAEGFAP 379

Db 356 QECVNTPGGFCECWGY-----EPGGP-----GEGACQVDDE-----CALGRSP- 395

QY 380 PHEPHRCQMFNCQACPADCDPNTQAS--CECPGEGYIL--DDGFICTDIDEC--ENGGFC 433

Db 396 -----CAQGC-TNTDGSFHCSEEGYVLAGEDGTCCQDVDECVGPGGPLC 439

QY 434 SGVCHNLPFTFECIGPDSALARIHIGTDCDSKV-----DGDGSGGE-----P 477

Db 440 DSLCFNTQSFHCGCLPGWILAPN-GVSCMTGMPVSLGPPSGPPDEEDKGEKSGSTVPRAA 498

QY 478 PPSPTPGSTLTTPPA 491

Db 499 TASPTRGEGTTPKA 512

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:44:08 ; Search time 16.5 Seconds  
(without alignments)  
2903.235 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516  
Perfect score: 2830  
Sequence: 1 APASPQGGSCVHDFAL.....PSPTFGSLTPPAVLVHSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78: \*  
1: Pir1: \*  
2: Pir2: \*  
3: Pir3: \*  
4: Pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID       | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 2826   | 99.9        | 575    | 1 THHUB  | thrombomodulin pre |
| 2          | 1858   | 65.7        | 577    | 2 A60501 | thrombomodulin pre |
| 3          | 1095.5 | 38.7        | 356    | 2 A25918 | thrombomodulin - b |
| 4          | 414    | 14.6        | 1574   | 2 T13954 | MEGF6 protein - ra |
| 5          | 414    | 14.6        | 1620   | 2 T27283 | hypothetical prote |
| 6          | 374    | 13.2        | 2907   | 2 A57278 | fibrillin-2 precu  |
| 7          | 368    | 13.0        | 2871   | 2 A55567 | fibrillin 1 - bovi |
| 8          | 367.5  | 13.0        | 1184   | 2 A55184 | fibulin-2 precurs  |
| 9          | 366    | 12.9        | 2918   | 2 A54105 | fibulin-2 precu    |
| 10         | 364    | 12.9        | 3002   | 2 A47221 | fibrillin 1 precu  |
| 11         | 362    | 12.8        | 2871   | 2 A55624 | fibrillin-1 precu  |
| 12         | 357    | 12.6        | 741    | 2 T46488 | hypothetical prote |
| 13         | 357    | 12.6        | 1221   | 2 A49457 | fibulin-2 precurs  |
| 14         | 332.5  | 11.7        | 1964   | 2 T09059 | notch4 - mouse     |
| 15         | 331    | 11.7        | 589    | 2 T43210 | fibulin-1D precurs |
| 16         | 330.5  | 11.7        | 1712   | 2 A38261 | masking protein pr |
| 17         | 324.5  | 11.5        | 689    | 2 T42760 | fibulin, splice fo |
| 18         | 324.5  | 11.5        | 712    | 2 T42990 | fibulin 1, splice  |
| 19         | 322.5  | 11.4        | 1394   | 2 A35626 | transforming growt |
| 20         | 321    | 11.3        | 2321   | 2 T78549 | notch3 protein - h |
| 21         | 314.5  | 11.1        | 3507   | 2 T34513 | hypothetical prote |
| 22         | 313    | 11.1        | 798    | 2 T22793 | hypothetical prote |
| 23         | 312.5  | 11.0        | 1820   | 2 A55494 | latent transformin |
| 24         | 311    | 11.0        | 601    | 2 B36346 | fibulin 1 precurs  |
| 25         | 311    | 11.0        | 683    | 2 C36346 | fibulin 1 precurs  |
| 26         | 309    | 10.9        | 685    | 2 S78040 | fibulin, splice fo |
| 27         | 308    | 10.9        | 705    | 2 S34968 | fibulin, splice fo |
| 28         | 303.5  | 10.7        | 1251   | 2 A57293 | latent transformin |
| 29         | 297    | 10.5        | 2531   | 2 A46019 | notch-1 protein -  |

ALIGNMENTS

RESULT 1

THHUB  
thrombomodulin precursor [validated] - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1988 #sequence revision 12-May-1995 #text change 15-Sep-2000  
C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954  
R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyama J. Biochem. 103, 281-285, 1988  
A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed activ  
A:Reference number: A41442; MUID:88227901; PMID:2836377  
A:Accession: A41442  
A:Molecule type: DNA  
A:Residues: 1-575 <SHL>  
A:Cross-references: DBJ:D00210; NID:G220126; PIDN:BAA00149.1; PID:G220127  
R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987  
A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the cDN  
A:Reference number: A28307; MUID:87317665; PMID:2819876  
A:Accession: A28307  
A:Molecule type: DNA; mRNA  
A:Residues: 1-472, 'A', 474-575 <JAC>  
A:Cross-references: GB:J02973; NID:G339658; PIDN:AAA61175.1; PID:G339659  
R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kawaha  
EMBO J. 6, 1891-1897, 1987  
A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endoth  
A:Reference number: A29680; MUID:88004395; PMID:2820710  
A:Accession: A29680  
A:Molecule type: mRNA  
A:Residues: 1-575 <SUZ>  
A:Cross-references: GB:X05495; NID:G37123; PIDN:CAA29045.1; PID:G736251  
A:Experimental source: lung endothelium  
A:Note: part of this sequence, including the amino end of the mature protein, were deter  
Biochemistry 26, 4350-4357, 1987  
A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of the  
A:Reference number: A27073; MUID:88024950; PMID:2822087  
A:Accession: A27073  
A:Molecule type: mRNA  
A:Residues: 1-472, 'A', 474-575 <WEN>  
A:Cross-references: GB:M16552; NID:G339656; PIDN:AAB59508.1; PID:G339657  
A:Experimental source: placenta  
A:Note: parts of this sequence were determined by protein sequencing  
R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.  
J. Biochem. 113, 433-440, 1993  
A:Title: Urinary thrombomodulin, its isolation and characterization.  
A:Reference number: JX0264; MUID:93293792; PMID:8390446  
A:Accession: JX0264  
A:Molecule type: protein; mRNA  
A:Residues: 19-472, 'A', 474-486 <YAM>  
A:Experimental source: urine  
A:Note: the urinary form appears to be identical with that circulating in plasma  
R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.O.; Grinnell, B.W.

cell-fate determin  
Notch homolog prot  
transmembrane prot  
notch protein - fr  
notch 3 protein -  
extracellular prot  
hypothetical prote  
epidermal growth f  
notch protein homo  
notch protein homo  
epidermal growth f  
Notch B protein -  
notch homolog - se  
Notch protein - Af  
growth potentiati  
growth arrest-spec

Biochem. J. 295, 131-140, 1993  
 A;Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.  
 A;Reference number: S38954; MUID:94029900; PMID:8216207  
 A;Accession: S38954  
 A;Molecule type: protein  
 A;Residues: 475-491, X, 493-494 <GER>  
 A;Note: the residue designated 'X' was determined to be a Ser with covalently bound chondroitin sulfate.  
 R;Meininger, D.P.; Komives, E.A.  
 submitted to the Brookhaven Protein Data Bank, September 1995  
 A;Reference number: A67369; PDB:1ZAO  
 A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residues 475-491  
 R;Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, August 1994  
 A;Reference number: A52804; PDB:1HLT  
 A;Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
 R;Hrabal, R.; Komives, E.A.; Ni, F.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A;Reference number: A65583; PDB:1FGD  
 A;Contents: annotation; conformation by (1)H-NMR, residues 427-444  
 R;Hrabal, R.; Komives, E.A.; Ni, F.  
 Protein Sci. 5, 195-203, 1996  
 A;Title: Structural resiliency of an EGF-like subdomain bound to its target protein, thrombospondin type 1.  
 A;Reference number: A58595; MUID:96276211; PMID:87453396  
 A;Contents: annotation; conformation by (1)H-NMR  
 C;Genetics:  
 A;Gene: GDB:THBD  
 A;Cross-references: GDB:119613; OMIM:188040  
 A;Map position: 20p11.2-20p11.2  
 A;Introns: #status absent  
 C;Complex: homodimer, urinary form  
 C;Function:  
 A;Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activation of plasminogen.  
 A;Pathway: blood coagulation/moderation  
 A;Note: the membrane-bound form is located on the endothelium luminal surface of arteries and is subject to endocytosis.  
 A;Note: thrombin complexed with the membrane-bound form is subject to endocytosis.  
 C;Superfamily: thrombospondin; C-type lectin homology; EGF homology  
 C;Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coagulation; protein  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-575/Product: thrombospondin, membrane-bound form #status predicted <MAT>  
 F;19-513/Domain: extracellular #status predicted <EXT>  
 F;19-486/Product: thrombospondin, urinary form #status experimental <MAU>  
 F;24-167/Domain: C-type lectin homology <LCH>  
 F;177-199/Region: PEST sequence  
 F;201-233/Region: PEST sequence  
 F;245-280/Domain: EGF homology <EG1>  
 F;288-323/Domain: EGF homology <EG2>  
 F;329-362/Domain: EGF homology <EG3>  
 F;369-404/Domain: EGF homology <EG4>  
 F;408-439/Domain: EGF homology <EG5>  
 F;445-480/Domain: EGF homology <EG6>  
 F;485-513/Region: PEST sequence  
 F;517-539/Domain: transmembrane #status predicted <TMN>  
 F;540-575/Domain: intracellular #status predicted <INT>  
 F;47, 115, 116, 382, 409/Binding site: carbohydrate (Aen) (covalent) #status predicted  
 F;174, 225, 411, 504/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F;245, 255, 252, 265, 267, 280, 288, 292, 308, 310, 323, 329, 340, 336, 349, 351, 362, 369, 378, 374-38  
 F;334, 498/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F;342/Modified site: erythro-beta-hydroxyasparagine (Aen) #status experimental  
 F;490, 492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental

Query Match 99.9%; Score 2826; DB 1; Length 575;  
 Best Local Similarity 99.8%; Pred. No. 2e-162;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 APAPPPQGGSCVHDCFPALPGPATFNLSQICDGLRGLHMTVRSSVAADVILLNGD 60  
 DB 19 APAPPPQGGSCVHDCFPALPGPATFNLSQICDGLRGLHMTVRSSVAADVILLNGD 78  
 QY 61 GGVGRRRLWIGLQPLPPGCGDKPKRLGFLRGFWMTGDNNTSYSRWALDLNGLPGLCV 120  
 DB 79 GGVGRRRLWIGLQPLPPGCGDKPKRLGFLRGFWMTGDNNTSYSRWALDLNGLPGLCV 138

QY 121 AVSAAEATVPSEPIWEQQCEVKADGFLCEHFFPATCRPLAVEPGAAAAAASITTYTTPA 180  
 DB 139 AVSAAEATVPSEPIWEQQCEVKADGFLCEHFFPATCRPLAVEPGAAAAAASITTYTTPA 198  
 QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGAVQGHWAREAPGAWDCSVENGCCHEACNA 240  
 DB 199 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGAVQGHWAREAPGAWDCSVENGCCHEACNA 258  
 QY 241 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCBHFCEHFFPATCRPLAVEPGAAAAAASITTYTTPA 300  
 DB 259 IFGAPRCQCPAGALQADGRSCTASATQSCNDLCBHFCEHFFPATCRPLAVEPGAAAAAASITTYTTPA 318  
 QY 301 DOHRCEDVDDCILEPSPORCVNTQGGFECHYCNLYDLDGECVPPVDPCEFRANCEYQC 360  
 DB 319 DOHRCEDVDDCILEPSPORCVNTQGGFECHYCNLYDLDGECVPPVDPCEFRANCEYQC 378  
 QY 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCQPCNCTACPADCDPNTQASCECEGYILLDDGFI 420  
 DB 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCQPCNCTACPADCDPNTQASCECEGYILLDDGFI 438  
 QY 421 CTIDECENGGFCGVCNHLPGTFCICGPDSSALAHHTGDCDCKVGDGSGSGEPSPS 480  
 DB 439 CTIDECENGGFCGVCNHLPGTFCICGPDSSALAHHTGDCDCKVGDGSGSGEPSPS 498  
 QY 481 PTFGSTLTTPPAVGLVHSG 498  
 DB 499 PTFGSTLTTPPAVGLVHSG 516

RESULT 2  
 A60501  
 thrombospondin precursor - mouse  
 N;Alternate names: fetomodulin  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text\_change 16-Jul-1999  
 C;Accession: S08488; A32001; A60501  
 R;Dittman, W.A.; Majerus, P.W.  
 Nucleic Acids Res. 17, 802, 1989  
 A;Title: Sequence of a cDNA for mouse thrombospondin and comparison of the predicted mouse  
 A;Reference number: S08488; MUID:89128454; PMID:2536925  
 A;Accession: S08488  
 A;Molecule type: mRNA  
 A;Residues: 1-577 <DIT>  
 A;Cross-references: EMBL:X14432; NID:g54781; PIDN:CAA32597.1; PID:g54782  
 R;Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
 J. Biol. Chem. 263, 15815-15822, 1988  
 A;Title: The structure and function of mouse thrombospondin. Phorbol myristate acetate st  
 A;Reference number: A32001; MUID:89008498; PMID:2844823  
 A;Accession: A32001  
 A;Molecule type: mRNA  
 A;Residues: 97-577 <DIT>  
 A;Cross-references: GB:J04060  
 R;Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwaseki, H.; Imada, M.  
 Dev. Biol. 140, 113-122, 1990  
 A;Title: Identification of fetomodulin, a surface marker protein of fetal development, as  
 A;Reference number: A60501; MUID:90292331; PMID:2162790  
 A;Accession: A60501  
 A;Molecule type: protein  
 A;Residues: 19-22;330-343;479-489;545-555;562-575 <IMA>  
 C;Comment: Thrombospondin binds to and internalizes with thrombin. It is also a cofactor  
 C;Superfamily: thrombospondin; C-type lectin homology; EGF homology  
 C;Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
 F;24-165/Domain: C-type lectin homology <LCH>  
 F;244-279/Domain: EGF homology <EG1>  
 F;287-322/Domain: EGF homology <EG2>  
 F;328-361/Domain: EGF homology <EG3>  
 F;368-403/Domain: EGF homology <EG4>  
 F;407-438/Domain: EGF homology <EG5>  
 F;444-479/Domain: EGF homology <EG6>

Query Match 65.7%; Score 1858; DB 2; Length 577;  
 Best Local Similarity 66.0%; Pred. No. 2.2e-104;







Db 1288 DGQCTNIFGEYRCLCYDGFMA5MDMKTCIDVNECDLNPNCMFCECENTKGSFICHQCL 1347  
 QY 336 NYDLVDGE--CVPFVDP--FRANCYQCOPLN-QTSYLCVCAEGFA-----PIP 380  
 Db 1348 GYSVKGGTGTCTD-VDECEIGHNCDMHASCLNVPVGSFKSCREGWVGNGIKICIDLECA 1406  
 QY 381 HEPHRCOMFCNQTACADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 430  
 Db 1407 NGTHQCSI-----NAQC-VNTPGSYRACSEGF-TGDGFTCSVDVDECAENLNCENGQ 1457  
 QY 431 -----GF-----CS-----GVCHNLPTPFCICGPDGALA 455  
 Db 1458 CLNVPYAYRCECEMGFTFASDRSCQDIDEC5FONICVFGTCNNLPGMFHICDDGVYELD 1517  
 QY 456 RHIG--TDGD 463  
 Db 1518 RTGNCCTDID 1527

RESULT 7

A55567

fibillin 1 - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 24-Nov-2003

C:Accession: A55567

R:Illustra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.

Genomics 23, 480-485, 1994

A:Title: Sequence of the coding region of the bovine fibillin cDNA and localization to

A:Reference number: A55567; MUID:95137597; PMID:7835900

A:Accession: A55567

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2871 <full>

A:Cross-references: GB:L28748; NID:9508427; PIDN:AAA74122.1; PID:9508428

C:Superfamily: fibillin; EGF homology

F:1201-1236/Domain: EGF homology <EGF>

Query Match 13.0%; Score 368; DB 2; Length 2871;  
 Best Local Similarity 31.5%; Pred. No. 1.6e-14;  
 Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 226 DCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP-- 282  
 Db 1200 ECSINWGGCTCTNSGSEYSCQCPGFALPDQASCT-----DIDE--CEDNPIC 1249

QY 283 -----DQGSYSYCMCTGYRLAADOHRCEVDVDDCILEPSPC-PQRCVNTQGFECVCP 335  
 Db 1250 DGQCTNIFGEYRCLCYDGFMA5MDMKTCIDVNECDLNPNCMFCECENTKGSFICHQCL 1309

QY 336 NYDLVDGE--CVPFVDP--FRANCYQCOPLN-QTSYLCVCAEGFA-----PIP 380  
 Db 1310 GYSVKGGTGTCTD-VDECEIGHNCDMHASCLNVPVGSFKSCREGWVGNGIKICIDLECA 1368

QY 381 HEPHRCOMFCNQTACADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 429  
 Db 1369 NGTHQCSI-----NAQC-VNTPGSYRACSEGF-TGDGFTCSVDVDECAENLNCENGQ 1419

QY 430 -----GF-----CS-----GVCHNLPTPFCICGPDGALA 455  
 Db 1420 CLNVPYAYRCECEMGFTFASDRSCQDIDEC5FONICVFGTCNNLPGMFHICDDGVYELD 1479

QY 456 RHIG--TDGD 463  
 Db 1480 RSGNCTDYNCELDPTTCISGNCVNTPGSYTCDPCPD-----FELNPRVGCV 1527

RESULT 8

A55184

fibulin-2 precursor - human

N:Alternate names: protein DKFP586A1519.1

C:Species: Homo sapiens (man)

C>Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 08-Sep-2002

A:Molecule type: DNA  
 A:Residues: 1-1620 <full>  
 A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y64G10A.f  
 A:Experimental source: clone Y64G10A  
 C:Genetics:  
 A:Gene: CESP:Y64G10A.f  
 A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7

Query Match 14.6%; Score 414; DB 2; Length 1620;  
 Best Local Similarity 27.0%; Pred. No. 1.8e-17;  
 Matches 128; Conservative 38; Mismatches 176; Indels 132; Gaps 20;

QY 101 YSRWRLDLNAPLCGPICVAVSAE-----ATVPSEPIWEBOCE--VKAGDFICEF 151  
 Db 56 YLRPAPFRRGCKCLLRQVANC5ADLCHNGTGVPSHNDNEQVCEPCVFTGAKCQY 115

QY 152 HPPATCRPLAVEPGAAAASVITYTTPFAARGADFOALPVSSAAVPLGLQMLCTAPP 211  
 Db 116 D-ANEC--MANNGGCEHCVN-TIGTY-----CRWPG 145

QY 212 AVQGHWAREAPGAWDCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSC- 270  
 Db 146 FELSDGNTCSIDISCAVNGGCSRCVNSPGFRCDCPSDLYLHADGRTG--GKVTSCS 203

QY 271 --NDLCEHFCVNPDPQSGYSYCMCTGYRLAADOHRCEVDVDDCILEPSPCPCPCVNTQGG 328  
 Db 204 TDNGGCEHC-ENDSNGEYRCKRVGFVXLSKSCQVDPFNDKGGCQCHCTNNHR 262

QY 329 FECHYVNDL-VDBECVFPVPCFRAN-CEYQCOPLNNTSYLCVCAEGFAPIPHEPHRC 386  
 Db 263 AOCQCYFGLSHYDRSCVDIDECANQNGEHCENVKGT-YRCKREGY-QIQRDGRTC 320

QY 387 QMF-----CNQTACPADC--DNPQASCECPGYYL----- 415  
 Db 321 EEMLGCGVNGGCGQDCHDQDPDGGHCKCRNGYILANDKQLCHDNI5VIHARAPRLWD 380

QY 416 -----DDGF-----ICTDIDEC-ENGFCGSGVC 437  
 Db 381 SYETVTCVTPTDLTCHKMLHDSGHVQCFDDGYELIDSKFCQDINECHENNGDCSQIC 440

QY 438 HNLPGFECICGPD5ALAHIGT-----DCDS-----GKVDGDSGSGEP 477  
 Db 441 VNLAGSVECOCKPFLMKDRKTCEDISECSNNGGCEQICSNQEGYWCSCPE 494

RESULT 6

A57278

fibillin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 24-Nov-2003

C:Accession: A57278

R:Zhang, H.; Hu, W.; Ramirez, F.

J. Cell Biol. 129, 1165-1176, 1995

A:Title: Developmental expression of fibillin genes suggests heterogeneity of extracell

A:Reference number: A57278; MUID:95263670; PMID:7744963

A:Accession: A57278

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2907 <ZHA>

A:Cross-references: GB:L39790; NID:g762830; PIDN:AAA74908.1; PID:g762831

C:Superfamily: fibillin; EGF homology

F:1239-1274/Domain: EGF homology <EGF1>

F:2488-2523/Domain: EGF homology <EGF>

Query Match 13.2%; Score 374; DB 2; Length 2907;  
 Best Local Similarity 32.9%; Pred. No. 7.2e-15;  
 Matches 102; Conservative 24; Mismatches 99; Indels 92; Gaps 18;

QY 226 DCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDP-- 283  
 Db 1238 ECMINWGGCTDCTNSGSEYSCSEGVAMPDGRSCA-----DIDE--CENNPIC 1287

QY 284 -----DQGSYSYCMCTGYRLAADOHRCEVDVDDCILEPSPC-PQRCVNTQGFECVCP 335

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A;Molecule type: mRNA  
A;Residues: 1-2918 <ZHA>  
A;Cross-references: GB:703272  
R;Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A;Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe  
A;Reference number: S17062; MUID:91304567; PMID:1852206  
A;Accession: S17063

A;Molecule type: mRNA  
A;Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A;Cross-references: EMBL:X62009  
R;Milewicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A;Reference number: S31101  
A;Accession: S31101

A;Molecule type: mRNA  
A;Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 192  
A;Cross-references: EMBL:X62009  
C;Genetics:  
A;Gene: GDB:FBN2  
A;Cross-references: GDB:128122; OMIM:121050  
A;Map position: 5q23-5q31  
C;Superfamily: fibrillin; EGF homology  
C;Keywords: extracellular protein  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-2918/Product: fibrillin-2 #status predicted <MAT>  
F;1245-1280/Domain: EGF homology <EGF1>  
F;1970-2013/Domain: EGF homology <EGF>

Query Match 12.9%; Score 366; DB 2; Length 2918;  
Best Local Similarity 33.1%; Pred. No. 2.2e-14;  
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;  
Qy 226 DCSVNGCGHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLCHEFCVNPDP-- 283  
Db 1244 ECMINNGGDTCTNSESSEYSCSEGVAMPDGRSCA-----DIDE--CENNPDIC 1293  
Qy 284 -----QPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPCP-QRCVNTGGFECCHCYP 335  
Db 1294 DGGQCTNIPGYRCLCYDGFWMASMDKTCIDVNECDLNSNCFMGECENTKGSFICHCQL 1353  
Qy 336 NYDLVDBG--CUEPVDPC--FRANCEYQCPQLN-QTSYLCVCAGGFADIPH-----EPHR 385  
Db 1354 GYSVKKGTGTCTD-VDECEIGHNCDMHASCLNIPGSKCSREGW--IGNGIKCIDLDE 1410  
Qy 386 COMFCNQACPADCPNTPQAS--CECPGYILDGFICTDIDE-----CENG----- 430  
Db 1411 CNGTHQCSINAOQ-VNTPGSVRCACSEGF-TGDGFTCSVDVDECAENLNCENGQCLNVP 1468  
Qy 431 -----GF-----CSGVCHNLPCTPECICGPDGALAHIG- 459  
Db 1469 GAYRCECEMGFTPASDRSCQDIBCSFQNICVSGTGNLPGMHFICICDDGYELDRGTGN 1528  
Qy 460 -TDGD 463  
Db 1529 CTDID 1533

RESULT 10  
A47221  
fibrillin 1 precursor - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 02-Jun-1995 #sequence revision 25-Apr-1997 #text change 24-Nov-2003  
C;Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198  
R;Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.  
Genomics 17, 476-484, 1993  
A;Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structu  
A;Reference number: A47221; MUID:94010947; PMID:7691719  
A;Accession: A47221  
A;Molecule type: mRNA  
A;Residues: 1-337, 'T', 339-1029 <COR>  
A;Cross-references: GB:X63556

C;Accession: A55184; T08744  
R;Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A;Title: Fibrillin-2 (FBN2): human cDNA sequence, mRNA expression, and mapping of the gen  
A;Reference number: A55184; MUID:95104855; PMID:7806230  
A;Accession: A55184  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1184 <ZHA>  
A;Cross-references: GB:X82494; NID:G575232; PIDN:CAA57876.1; PID:G575233  
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A;Reference number: Z16471  
A;Accession: T08744

A;Molecule type: mRNA  
A;Residues: 656-719, 'QDECLMGHDCRRQFCVNTLGSFYCVNHTVLCADGYILNAHRKCV', 720-853, 'T', 855-11  
A;Cross-references: EMBL:AL050095  
A;Experimental source: adult uterus; clone DKFZp586A1519  
C;Genetics:  
A;Gene: GDB:FBN2  
A;Cross-references: GDB:293037; OMIM:135821  
A;Map position: 3p25-3p24  
A;Note: DKFZp586A1519.1  
C;Superfamily: fibrillin-2; EGF homology; von Willebrand factor type C repeat homology  
C;Keywords: alternative splicing; extracellular matrix  
F;1-27/Domain: signal sequence #status predicted <SIG>  
F;28-1184/Product: fibrillin-2 protein #status predicted <MAT>  
F;905-941/Domain: EGF homology <EGF>

Query Match 13.0%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 8.6e-15;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
Qy 136 BEQCEVKAD-GFCEHFPATC---RPLAV-----EPGAAAASVITGTPFAARG 183  
Db 531 EGQCESNPNI/GYPCN-HVMSCEGEPLVPEVRPPPEAARRVS-----EAEM 582  
Qy 184 ADFQALPVGSAAV---APLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCGEACNA 240  
Db 593 AGREALSGTAEALPNSLPGDDQDECLLPBEL-----COHLCLIN 622  
Qy 241 IPGAPRCQCPAGALQADGRSC-----TASAT-----QS 269  
Db 623 TVGSHYKACFPGLSLQDDGTCRPEGHPQPEAPPEALKSFQSVANITPLPQNT 682  
Qy 270 QND--LCEHFCVNPDPQPSYSCMCTGYRLAADQHRCEVDVDCILEPSPCP--QRCVNT 325  
Db 683 CKDNGPCKQVC---STVGGSAICSEFPFYAIMADGVSCEDINECVTDLHTCSRGEHCNT 739  
Qy 326 QGGFECH---CVPNYDLVDGCEVPEVDFRANCEYQCPQLNQTSLCV----- 371  
Db 740 LGSFHYKALTCEFGYALKDGE-EDVDEC--AMGHTTQCP---GFLCNTKGSFYCOA 792  
Qy 372 ---CAEGFAPIPH-----EPHRCMFNQATCAPDCDNTQASCECPGY-I 414  
Db 793 RQRCDMGFLQDPGNCVNDINECTSLSEPCRPFGSCINTVGSYTCQRNPLI---CARGYHA 849  
Qy 415 LDGFICTDDECEGFCGSG---VCHNLPCTPECIC 448  
Db 850 SDDGAKCVDNCEGTGVRHCGEGVCHNLPGSYRCD 886

RESULT 9  
A54105  
fibrillin-2 precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 09-Sep-1994 #sequence revision 09-Sep-1994 #text change 24-Nov-2003  
C;Accession: A54105; S17063; S31101  
R;Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham,  
J. Cell Biol. 124, 855-863, 1994  
A;Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe  
A;Reference number: A54105; MUID:94165150; PMID:8120105  
A;Accession: A54105

R;Pereira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglilan, T.; Bonadum. Mol. Genet. 2, 961-968, 1993  
A;Title: Genomic organization of the sequence coding for fibrillin, the defective gene in Marfan syndrome.  
A;Reference number: 154355; MUID:93372860; PMID:8364578  
A;Accession: I54355  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 132-3002 <PER>  
A;Cross-references: GB:I13923; NID:g306745; PIDN:AAB02036.1; PID:g306746  
A;Cross-references: GB:S54426; NID:g264860; PIDN:AAB25244.1; PID:g264861  
R;Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991  
A;Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A;Reference number: S17064; MUID:91304568; PMID:1852207  
A;Accession: S17064  
A;Molecule type: mRNA  
A;Residues: 1030-3002 <MAS>  
A;Cross-references: EMBL:X63556  
R;Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A;Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A;Reference number: 159574; MUID:93157831; PMID:8430317  
A;Accession: 159574  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 2217-2288, '1', 2290-2325 <RES>  
A;Cross-references: GB:S54426; NID:g264860; PIDN:AAB25244.1; PID:g264861  
R;Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattel, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A;Title: Linkage of Marfan syndrome and a phenotypically related disorder to two different chromosomes.  
A;Reference number: S17062; MUID:91304567; PMID:1852206  
A;Accession: S17062  
A;Molecule type: mRNA  
A;Residues: 'VLTVVFILSYNKM', 944-1444 <LEE1>  
A;Cross-references: EMBL:X62008; NID:g31398; PIDN:CAB56534.1; PID:g5924015  
A;Accession: S62111  
A;Molecule type: protein  
A;Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>  
R;Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A;Title: Connective tissue microfibrils. Isolation and characterization of three large R  
A;Reference number: A34198; MUID:90078246; PMID:2512293  
A;Accession: A34198  
A;Molecule type: protein  
A;Residues: 565-575; 1890-1892, 'I', 1894-1900 <MAD>  
C;Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C;Genetics:  
A;Gene: GDB:FBN1  
A;Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A;Map position: 15q21.1-15q21.1  
A;Introns: 2236/1; 2258/1; 2297/1  
C;Superfamily: fibrillin; EGF homology  
C;Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; F;1-3002/Product: fibrillin (5'-region exon A splice form) #status predicted <MAD>  
F;1332-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MAD>  
F;1332-1367/Domain: EGF homology <EGF>  
F;1457-1492/Domain: EGF homology <EGF2>  
F;2262-2295/Domain: EGF homology <EGF1>

381 HEPHRCQFNCQTACPADCDNTQAS--CECEGVYLDGFICTDIDEC-EN----- 429  
1500 NGTHWCSSOH-----ADC-KNTWGSYRCLCKEY--TGDGFTCTDLDECSENLCNGQ 1550  
430 -----GGF-----CS-----GVCHNLPGTFECICGPDLSALA 455  
1551 CLNAPGGYRCECDMGFVPSADGKACEDIDECLEPNICVFGTCHNLPGFRCCECEGYELD 1610  
456 RHIG-----TDCDSKGVDDSGSGEPSPPTGS-----TLTPPAVGLV 495  
1611 RSGNGCTDNECLDPTTCISGNCVN-----TPGSYICDPPDFELNPTRVGCV 1658  
RESULT 11  
A;Accession: A55624  
A;Molecule type: mRNA  
A;Residues: 1-2871 <VIN>  
A;Cross-references: GB:L29454; NID:g575509; PIDN:AAA56840.1; PID:g575510  
C;Genetics:  
A;Gene: Fbn-1  
C;Superfamily: fibrillin; EGF homology  
F;1201-1236/Domain: EGF homology <EGF>  
Query Match 12.8%; Score 362; DB 2; Length 2871;  
Best Local Similarity 29.4%; Pred. No. 3.7e-14;  
Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;  
194 SAAVAPLGL-----QLMCTAPPAGVQGHVAREAPAGWDCSVENGCGCHACNAIPGAPRC 247  
1162 SANLPHRCVNLIGKYQACNPGVPHTRDLFCVDIDECSIMNGGCTFCTNSDGSYEC 1221  
248 QCPAALQADGRSTASATOSCNLDCEHFCVPNP-----DQGSYSQMCETGYRL 298  
1222 SCQPGFALMPDQSRCT-----DIDQ--CEDNPNICDGGCTNIPREYRCLCYDGFNA 1271  
299 ADQHRCEVDVDCILEPSPC-PQRCVNTQGGPECHYFNIDLVGB--CPEVDPC--FR 353  
1272 SEDMTCTVDVNECDLNPICLSGTCENTKGSFICHCDMGYSKKGKTGCTD-INECEIGA 1330  
354 ANCEYQCOPLNOT-SYLCVCAEGFA-----PIPEHRCQFNCQTACPADCDN 402  
1331 HNGRHAVENTAGSPKSCSPGWIGDGIKTDLDECSTHMCQSH-----ADC-KN 1382  
403 TQAS--CECEGVYLDGFICTDIDEC-EN-----GGF----- 432  
1383 TMGSYRCLCKOGY--TGDGFTCTDLDECSENLCNGQCLNAPGGYRCECDMGFVPSADG 1441  
433 -----CS-----GVCHNLPGTFECICGPDLSALRHIG-----TDCDSK 466  
1442 KACEDIDECLEPNICVFGTCHNLPGFRCCECEGYELDRSGGCTDNECLDPTTCISG 1501  
467 VDGSDSGSGEPSPPTGS-----TLTPPAVGLV 495  
1502 CVN-----TPGSYTCDCSPDFELNPTRVGCV 1527  
RESULT 12  
T46488  
hypothetical protein DKFZp434J065.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C;Accession: T46488

Query Match 12.9%; Score 364; DB 2; Length 3002;  
Best Local Similarity 30.8%; Pred. No. 2.9e-14;  
Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;  
226 DCSVENGGCEHACNAIPGAPRCQCPAGALQADGRSTASATQSCNLDCEHFCVPNP--- 282  
1331 ECSIMNGGCEFTCTNSGSEYSCSPGFPALMPDQSRCT-----DIDE--CEDNPNIC 1380  
283 -----DQGSYSQMCETGYRLADQHRCEVDVDCILEPSPC-PQRCVNTQGGPECHCYP 335  
1381 DGGQCTNIPGEYRCLCYDGFNAESDMKTCVDVNECDLNPICLSGTCENTKGSFICHDM 1440  
336 NYDLVDGE--CPEVDPC--FRANCEYQCOPLNOT-SYLCVCAEGFA-----PIP 380  
1441 GYSKKGKGTGCTD-INECEIGHNCGKHAVCTNTAGSPKSCSPGWIGDGIKTDLDEC 1499

Wed Jun 9 13:16:02 2004

R,Dueterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23035  
A:Accession: T46488  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-741 <AAA>  
A:Cross-references: EMBL:AL137638  
A:Experimental source: adult testis; clone DKFZp434J065  
C:Genetics:  
A:Note: DKFZp434J065.1

Query Match 12.6%; Score 357; DB 2; Length 741;

Best Local Similarity 31.6%; Pred. No. 2.5e-14;  
Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;  
227 CSVNGGCEHACNAIPGAPROCPAGALQADGRCTA-SATQSCNDLCEHFCVNPDPQP 285  
124 CAMEHDNCEBOLCVNVPVSGFVCCQYSGYALAEKGRCAVDYCAENHGCEHCV-NAD-- 124  
286 GSYSCMCEGYRLAADOHRCEDVDDCILPEPCPCPCVNTQGGFECHCYPNYDL-VDGEC 344  
125 GSYLCQCHGEGFALNPDEKTKIDYCASNHGCEHCVNTDSDYSCHLKFGLNPKKT 184  
345 VEPVDFCF-RANCEYQCCPLNTSYLVCACGAFPIPH-----EPHRCQMEC 390  
185 CRRINYCALNKPGEHCVNMEB-SYYCRCHRGVTLDPNGKTCRVHDHCAQDHGCEQLC 243  
391 NOTACPADCPNTQAS--CSCPGYILDDGF-ICTDIDEC---ENGFGCGVCHNLPGET 444  
244 -----LNTEDSVFVCCSEGLNEDLKTCRVYDCLLSHDG--CEYSCVNMDRSF 291  
445 ECICGPPSALARIHTDCCDSKVDG---GDSG 473  
292 ACQC-PGHLVRSQKTC--AKLDSALGDHG 320

RESULT 13

A49457  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 08-Sep-2002  
C:Accession: A49457; S74095  
R:Pas, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.  
J. Cell Biol. 123, 1269-1277, 1993  
A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with  
A:Reference number: A49457; MUID:94064787; PMID:8245130  
A:Accession: A49457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1221 <PAN>  
A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047  
R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.  
Eur. J. Biochem. 240, 427-434, 1996  
A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases  
A:Reference number: S74094; MUID:96439073; PMID:8841408  
A:Accession: S74095  
A:Molecule type: protein  
A:Residues: 236-238, X', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-589  
C:Superfamily: fibulin-2; EGF homology; von Willebrand factor type C repeat homology  
C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer  
F:942-978/Domain: EGF homology <EGF>

Query Match 12.6%; Score 357; DB 2; Length 1221;  
Best Local Similarity 29.5%; Pred. No. 3.7e-14;  
Matches 123; Conservative 33; Mismatches 141; Indels 114; Gaps 27;  
136 EEOQCEVKAD-GFLCEHFPATC-----RPLAV-----EFGAAAASV-----IT 174  
521 EGQSCENPNLGYFCN-HVMSLSCGEGEPLVPRVRRPPEAPRRVSEMEMASREALS 579  
175 YGT-----PFAARGADFO---ALP-----VGSSAAVAPGLQLM-----CTAPP 211

580 LGTEAELENSLPDGDQDECLMLPGLCOHLICINTVSGYRCACFPGLQGDGRTCPDRG 639  
212 AVQGHWARE-APGAWDCSV-----ENGGEHACNAIPGAPROCPAGAA 254  
640 APQDITAREAPRSESQVSPNTIPLVPQPNCKDNGPCQVCRVVGDTAMCSCFFGYA 699  
255 LQADGRSC-----TASATQSCNDLCEHFCVNPDPQGSYS- -MCETGYRLAADOHR 304  
700 IMADGVSCEDQDECLMGTHDCS--WKQFCV---NTLGSFYCVNHTVLCBEGYILNA-HRK 753  
305 CEDVDDCILPESPC--PQRCVNTQGGFECH---CYPNYDLVDGCEVPEVDFCFRA--NC 356  
754 CVDINECVTDLTCTRAEHCVNTPGSPCYKALICEPGYVLTGECTD-VDECVTGTTHNC 812  
357 E--YOCQPLNTSYLVC---CARGFAPIPH-----EPHRCQMECQNTACPA 397  
813 QAGFSQCN-TKGSFYCOARQRCMDGFLQDPGECNVINECTSLLEPCRSFSGSCINTVGSY 871  
398 DCDPNTQASCCEPGY-ILDDGFICTDIDECENGFGCSG---VCHNLPGTFECICGP 450  
872 TCORNPLV---CGRGYHANESEGCVDVNECETGVHRCGEGQLCYNLPDGYRCDCPK 925  
RESULT 14  
T09059  
notch4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 08-Sep-2002  
C:Accession: T09059  
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.B.; Dankers, C.; Lasky, C.; Sc  
submitted to the EMBL Data Library, October 1997  
A:Description: Sequence of the mouse major histocompatibility locus class III region.  
A:Reference number: Z16543  
A:Accession: T09059  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1964 <ROW>  
A:Cross-references: EMBL:AF030001; NID:92564945; PID:92564947  
C:Genetics:  
A:Gene: notch4  
A:Map position: 17  
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67  
1679/3; 1729/1; 1761/3  
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology  
C:Keywords: receptor; signal transduction  
F:514-545/Domain: EGF homology <EGF>  
Query Match 11.7%; Score 332.5; DB 2; Length 1964;  
Best Local Similarity 26.3%; Pred. No. 1.6e-12;  
Matches 128; Conservative 34; Mismatches 142; Indels 183; Gaps 31;  
116 GPLCVAVSAAATVPSEPIWEEQQCEVKADGFLCB-FHFPATCRPLAVEPGAAAASVIT 174  
35 GGTCLRLRGGGIC-----QC---APGLGETCQFPDPCR----- 66  
175 YGTPFAARGADFOAL---PVGSSAAVAPGLQLMCTAPG-----AVQGHWAREAPGAWDC 227  
67 -DTQLCKNGSCQALLPTPPSRSRPTSLTPHFSTCTSPSGFTGDRCTHLELCLCPSPF-C 124  
228 SYVNGGCEHACNAIPGAPROCPAGALQADGRCTASATQSCNDLCEHFCVNPDPQP- 286  
125 S--NGG--HCYVOAGSRPOCCSEPGWT---GEQCC-----LRDFCSANPCANGG 166  
287 -----SYSCMCEGYRLAADOHRC-VDVDCILPESPCPQ--RCVNTQGGFECH- 333  
167 VCLATVPQICRCPFGF-----EGHTCERDINECFLEPPGCPQGTSCNTLGSYQCLCPVG 222  
334 --YPNYDLVDGCEVPEVDFCFRAVCEYQCPPL---NQTSLVCVCAEGFAPIPHE----- 382  
223 QEGPOCKLRKAGC--PPGSLNGG---TCOLVPEGHSTHLCCLCPPGFTGLDCMNPDDC 277  
383 -PHRCQMECQNTACPADCDPNTQASCCEPGYILDDGFICT-DIDECE-----NGGF 432

Wed Jun 9 13:16:02 2004

278 VRHQQ---NGATCLDGLDVT---CLCPKTW---KGWDCSEIDICEARGPPRCRNGGT 328  
433 C-----SGVCHNLPGTFECICGP-----450  
329 QNTAGSPHCVCVSGWGAGCEENLDDCAAAATCAPGSTCIDRVGSFSLCLCPPGRTGLLCH 388  
451 --DSALAR--HIGTDCDSKVDG-----GDSGS-----GEPPSP-----481  
389 LEDMCLSPCHVNAQCSNPLTGSPLCICQPGYSGSTCHQDLDECQMAQOGSPCEHGS 448  
482 ---TEGS 485  
449 CINTFGS 455  
Db

RESULT 15  
T43210  
fibulin-1D precursor - Caenorhabditis elegans (fragment)  
C:Species: Caenorhabditis elegans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 08-Sep-2002  
C:Accession: T43210  
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.  
submitted to the EMBL Data Library, June 1998  
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character  
A:Reference number: Z22337  
A:Accession: T43210  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-589 <BAR>  
A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1  
C:Genetics:  
A:Note: intron positions not resolved (incomplete sequence)  
C:Superfamily: fibulin-1; EGF homology

Query Match 11.7%; Score 331; DB 2; Length 589;  
Best Local Similarity 28.1%; Pred. No. 7.5e-13;  
Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;  
QY 85 GPLRGP-QWVTGDNNTSYSRWARLDLNGAPLCGLCVASAAEATVPSEPI-----WEEQ 138  
Db 19 GCLRSFNKCCNGDIETH---ASEIITGRPLNDPHVLHLGDRKASSHCEHLCHDRGGKEV 75  
QY 139 QCEVKA-----DGFCLCEHFPF--ATCRPLAVE-----PGAAAAVSIITYGTPEA--- 180  
Db 76 ECSRSRSGFDLAPDGNACVDHIDECATLMDCLQSQRCLNTPGSKCIRTLSCGTGYAMDS 135  
QY 181 --ARGADFOALPVGSSAAVAPLGLQIMCTAPPVAVQ-----GHWAREAPGAWDCS---- 228  
Db 136 ETERCRDVDECNLGSN-----DCGPLYQCRNTQGSYRCDAKKKGDELQNPMTGECTSITC 191  
QY 229 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGAAALQADGSC 264  
Db 192 PNGYYPKNGMCMNDIDECVTGHNCGAGEECVNTPGSRFCQCKGNLCAHYEVN----- 243  
QY 265 SATQSCNDL--CEH-----FCVPNDPQGSYSCMCTGYRLAADQHRCEDDVDDCIL---- 313  
Db 244 GATGFCEDVNECQQGVCGSMCEI---NLFGTYKCKCPGPFENDAKKRCEDVDECIKFAG 300  
QY 314 EPSFCPQRQVNTGSGFECCHYPNYDLV--DGECEVPVDPQFR--ANCEYQCQPLNOTSILC 370  
Db 301 HVCDSLAEICINTIGSFCECKPGFOLASDGRERDNECTTGTAACEQKCVNI--PGSYQC 359  
QY 371 VCAEGFAPIP-----HEPHRCQMF--NOTACPADCDPNTQAS--CECPGY-ILDDGFI 420  
Db 360 ICDRGFALGPDGDKCEDIDECISIWAGSNDLCMGGC--INTKGYLQCPFGYKIQPDRT 418  
QY 421 CTDIDECENGFGCSG-----VCHNLPGTFEC 446  
Db 419 CVDVDECAMGE-CAGSDKVCVNTLGSFKC 446

Wed Jun 9 13:16:02 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:42:38 ; Search time 10.5 Seconds  
(without alignments)  
2469.613 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516

Perfect score: 2830  
Sequence: 1 APAEPQGGSCVHDFCAL.....PSPTFGSTLTPPAVGLVHSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------|
| 1          | 2830   | 100.0       | 575    | 1  | TRBM_HUMAN  |
| 2          | 1858   | 65.7        | 577    | 1  | TRBM_MOUSE  |
| 3          | 1095.5 | 38.7        | 356    | 1  | TRBM_BOVIN  |
| 4          | 582    | 20.6        | 652    | 1  | CD93_HUMAN  |
| 5          | 542    | 19.2        | 644    | 1  | CD93_MOUSE  |
| 6          | 525.5  | 18.6        | 643    | 1  | CD93_RAT    |
| 7          | 374    | 13.2        | 2907   | 1  | FN2_MOUSE   |
| 8          | 367.5  | 13.0        | 2871   | 1  | FN1_BOVIN   |
| 9          | 367.5  | 13.0        | 1184   | 1  | FN2_HUMAN   |
| 10         | 366    | 12.9        | 2911   | 1  | FN1_PIG     |
| 11         | 366    | 12.9        | 2871   | 1  | FN2_HUMAN   |
| 12         | 364    | 12.9        | 2871   | 1  | FN1_MOUSE   |
| 13         | 362    | 12.8        | 2871   | 1  | FN1_MOUSE   |
| 14         | 357    | 12.6        | 1221   | 1  | FN2_MOUSE   |
| 15         | 356    | 12.6        | 956    | 1  | MTN2_HUMAN  |
| 16         | 332.5  | 11.7        | 1964   | 1  | NTC4_MOUSE  |
| 17         | 331.5  | 11.7        | 1389   | 1  | LTBL_MOUSE  |
| 18         | 331.5  | 11.7        | 1713   | 1  | LTBL_RAT    |
| 19         | 330.5  | 11.7        | 1712   | 1  | MTN2_MOUSE  |
| 20         | 326.5  | 11.5        | 956    | 1  | LTBS_HUMAN  |
| 21         | 322.5  | 11.4        | 1394   | 1  | NTC3_HUMAN  |
| 22         | 322.5  | 11.4        | 1595   | 1  | NTC3_HUMAN  |
| 23         | 321    | 11.3        | 2321   | 1  | NTC4_HUMAN  |
| 24         | 315.5  | 11.1        | 443    | 1  | FB1_MOUSE   |
| 25         | 313    | 11.1        | 798    | 1  | FB1_MOUSE   |
| 26         | 311    | 11.0        | 703    | 1  | FB1_MOUSE   |
| 27         | 308.5  | 10.9        | 443    | 1  | FB1_MOUSE   |
| 28         | 308.5  | 10.9        | 2003   | 1  | NTC4_HUMAN  |
| 29         | 308    | 10.9        | 448    | 1  | FB1_MOUSE   |
| 30         | 307    | 10.8        | 704    | 1  | FB1_MOUSE   |
| 31         | 306.5  | 10.8        | 443    | 1  | FB1_MOUSE   |
| 32         | 306    | 10.8        | 681    | 1  | FB1_MOUSE   |
| 33         | 306    | 10.8        | 705    | 1  | FB1_MOUSE   |

|    |       |      |      |   |            |
|----|-------|------|------|---|------------|
| 34 | 302   | 10.7 | 448  | 1 | FB1_MOUSE  |
| 35 | 297   | 10.5 | 2531 | 1 | NTC1_MOUSE |
| 36 | 296   | 10.5 | 448  | 1 | FB1_MOUSE  |
| 37 | 294   | 10.4 | 2319 | 1 | NTC3_RAT   |
| 38 | 292   | 10.3 | 2470 | 1 | NTC2_MOUSE |
| 39 | 291   | 10.3 | 598  | 1 | FB1_MOUSE  |
| 40 | 289.5 | 10.2 | 493  | 1 | FB1_MOUSE  |
| 41 | 287   | 10.1 | 2471 | 1 | NTC2_RAT   |
| 42 | 285   | 10.1 | 2471 | 1 | NTC2_HUMAN |
| 43 | 284.5 | 10.1 | 2437 | 1 | NTC1_MOUSE |
| 44 | 280.5 | 9.9  | 2703 | 1 | NTC1_MOUSE |
| 45 | 280   | 9.9  | 2318 | 1 | NTC3_MOUSE |

## ALIGNMENTS

| RESULT 1 | TRBM_HUMAN   | STANDARD | PRT | 575 AA. |
|----------|--|----------|-----|---------|
| ID       | P07204   |          |     |         |
| DT       | 01-APR-1988 (Rel. 07, Created)   |          |     |         |
| DT       | 01-FEB-1991 (Rel. 17, Last sequence update)                            |          |     |         |
| DT       | 15-MAR-2004 (Rel. 43, Last annotation update)                          |          |     |         |
| DE       | Thrombomodulin precursor (Petomodulin) (TM) (CD141 antigen).           |          |     |         |
| GN       | THBD OR THRM.  |          |     |         |
| OS       | Homo sapiens (Human).  |          |     |         |
| OC       | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |          |     |         |
| OC       | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.             |          |     |         |
| OX       | NCBI_TaxID=9606;   |          |     |         |
| RN       | [1]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | MEDLINE=88004395; PubMed=2820710;                                      |          |     |         |
| RA       | Suzuki K., Kusumoto H., Deyashiki Y., Nishioka J., Maruyama I.,        |          |     |         |
| RA       | Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.,            |          |     |         |
| RT       | "Structure and expression of human thrombomodulin, a thrombin          |          |     |         |
| RT       | receptor on endothelium acting as a cofactor for protein C             |          |     |         |
| RT       | activation.";  |          |     |         |
| RL       | EMBO J. 6:1891-1897(1987).   |          |     |         |
| RN       | [2]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | MEDLINE=88024950; PubMed=2822087;                                      |          |     |         |
| RA       | Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E., |          |     |         |
| RT       | "Human thrombomodulin: complete cDNA sequence and chromosome           |          |     |         |
| RT       | localization of the gene.";  |          |     |         |
| RL       | Biochemistry 26:4350-4357(1987).                                       |          |     |         |
| RN       | [3]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | MEDLINE=87317665; PubMed=2819876;                                      |          |     |         |
| RA       | Jackman R.W., Beiler D.L., Fritze L., Soff G., Rosenberg R.D.,         |          |     |         |
| RT       | "Human thrombomodulin gene is intron depleted: nucleic acid sequences  |          |     |         |
| RT       | of the cDNA and gene predict protein structure and suggest sites of    |          |     |         |
| RT       | regulatory control.";  |          |     |         |
| RL       | Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).                      |          |     |         |
| RN       | [4]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | MEDLINE=88227901; PubMed=2836377;                                      |          |     |         |
| RA       | Shirai T., Shoji S., Ito H., Yamamoto S., Kusumoto H.,                 |          |     |         |
| RA       | Deyashiki Y., Maruyama I., Suzuki K.,                                  |          |     |         |
| RT       | "Gene structure of human thrombomodulin, a cofactor for thrombin-      |          |     |         |
| RT       | catalyzed activation of protein C.";                                   |          |     |         |
| RL       | J. Biochem. 103:281-285(1988).   |          |     |         |
| RN       | [5]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,       |          |     |         |
| RA       | Ozuna M., Roel C.L., Toth E.J., Yi Q., Nickerson D.A.;                 |          |     |         |
| RA       | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.                |          |     |         |
| RN       | [6]  |          |     |         |
| RP       | SEQUENCE FROM N.A.   |          |     |         |
| RX       | MEDLINE=21638749; PubMed=11780052;                                     |          |     |         |
| RA       | Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,     |          |     |         |
| RA       | Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,     |          |     |         |
| RA       | Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,            |          |     |         |

- RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clapp M.E., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
RA Ellington A.G., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.R., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.L., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., King A., Knights K., Laird G.K., Lawlor S.,  
RA Leivaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Peck A.I.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RA "The DNA sequence and comparative analysis of human chromosome 20.";  
RA Nature 414:865-871(2001).  
RA [7]  
RA SEQUENCE FROM N.A.  
RA TISSUE=Lung;  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano J.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krywinski M.I., Skalska U., Smalus D.B.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RA human and mouse cDNA sequences";  
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RA [8]  
RA CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RA MEDLINE=94029900; PubMed=8216207;  
RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,  
RA Grinnell B.W.;  
RA "Identification of the predominant glycosaminoglycan-attachment site  
RA in soluble recombinant human thrombomodulin: potential regulation of  
RA functionality by glycosyltransferase competition for serine474";  
RA Biochem. J. 295:131-140(1993).  
RA [9]  
RA STRUCTURE BY NMR OF 389-407.  
RA MEDLINE=96007474; PubMed=7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RA "The structure of a 19-residue fragment from the C-loop of the fourth  
RA epidermal growth factor-like domain of thrombomodulin";  
RA J. Biol. Chem. 270:23366-23372(1995).  
RA [10]  
RA STRUCTURE BY NMR OF 364-407.  
RA MEDLINE=96100636; PubMed=8528067;  
RA Meininger D.P., Hunter M.J., Komives E.A.;  
RA "Synthesis, activity, and preliminary structure of the fourth  
RA EGF-like domain of thrombomodulin";  
RA Protein Sci. 4:1683-1695(1995).  
RA [11]  
RA STRUCTURE BY NMR OF 427-444.  
RX MEDLINE=95034791; PubMed=7947766;  
RX Sinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RA "Thrombin-bound structure of an EGF subdomain from human  
RA thrombomodulin determined by transferred nuclear Overhauser  
RA effects";  
RL Biochemistry 33:13553-13560(1994).  
RN [12]  
RP STRUCTURE BY NMR OF 427-444.  
RP MEDLINE=96276211; PubMed=8745396;  
RP Hrabal R., Komives E.A., Ni F.;  
RA "Structural resiliency of an EGF-like subdomain bound to its target  
RA protein, thrombin";  
RL Protein Sci. 5:195-203(1996).  
RN [13]  
RP STRUCTURE BY NMR OF 405-444.  
RP MEDLINE=98035729; PubMed=9367781;  
RP Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;  
RA "Structure of the fifth EGF-like domain of thrombomodulin: an  
RA EGF-like domain with a novel disulfide-bonding pattern";  
RL J. Mol. Biol. 273:913-926(1997).  
RN [14]  
RP VARIANT TED TYR-486.  
RP MEDLINE=95111115; PubMed=7811989;  
RP Oehlin A.-K., Marlar R.A.;  
RA "The first mutation identified in the thrombomodulin gene in a  
RA 45-year-old man presenting with thromboembolic disease";  
RL Blood 85:330-336(1995).  
RN [15]  
RP VARIANT TED TYR-486, AND VARIANTS THR-43; ALA-79; SER-495 AND LEU-501.  
RP MEDLINE=97341986; PubMed=9198186;  
RP Oehlin A.-K., Norlund L., Marlar R.A.;  
RA "Thrombomodulin gene variations and thromboembolic disease";  
RL Thromb. Haemost. 78:396-400(1997).  
RN [16]  
RP VARIANT VAL-473.  
RP MEDLINE=97206518; PubMed=9157575;  
RP Norlund L., Holm J., Zoller B., Oehlin A.-K.;  
RA "A common thrombomodulin amino acid dimorphism is associated with  
RA myocardial infarction";  
RL Thromb. Haemost. 77:248-251(1997).  
RN [17]  
RP VARIANT THR-43.  
RP MEDLINE=99057299; PubMed=9843165;  
RP Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stoggen P.J., Manger Cats V., Ireland H.;  
RA "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
RA and the risk of myocardial infarction in men";  
RL Thromb. Haemost. 80:743-748(1998).  
RN [18]  
RP VARIANT VAL-473.  
RP MEDLINE=21143723; PubMed=11245641;  
RP Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RA "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
RA disease";  
RL Circulation 103:1386-1389(2001).  
RN [19]  
RP VARIANT TED TYR-486, AND VARIANT VAL-473.  
RP MEDLINE=22135346; PubMed=12139752;  
RP Faioni E.M., Franchi F., Castaman G., Biguzzi E., Rodeghiero F.;  
RA "Mutations in the thrombomodulin gene are rare in patients with  
RA severe thrombophilia";  
RL Br. J. Haematol. 118:595-599(2002).  
CC -!- FUNCTION: Thrombomodulin is a specific endothelial cell receptor  
CC that forms a 1:1 stoichiometric complex with thrombin. This  
CC complex is responsible for the conversion of protein C to the  
CC activated protein C (protein Ca). Once evolved, protein Ca  
CC scissions the activated cofactors of the coagulation mechanism,  
CC factor Va and factor VIIIa, and thereby reduces the amount of  
CC thrombin generated.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing  
CC thrombomodulin.



|                                 |   |        |     |
|---------------------------------|---|--------|-----|
| us-09-509-994-2_copy_19_516.rsp |   | Page 3 |     |
| CC                              | activated protein C (protein Ca). Once evolved, protein Ca  |        |     |
| CC                              | scissions the activated cofactors of the coagulation mechanism,   |        |     |
| CC                              | factor Va and factor VIIIa, and thereby reduces the amount of   |        |     |
| CC                              | thrombin generated.   |        |     |
| CC                              | !- SUBCELLULAR LOCATION: Type I membrane protein.   |        |     |
| CC                              | !- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing   |        |     |
| CC                              | thrombomodulin (by similarity).   |        |     |
| CC                              | !- SIMILARITY: Contains 6 EGF-like domains.   |        |     |
| CC                              | This SWISS-PROT entry is copyright. It is produced through a collaboration                                      |        |     |
| CC                              | between the Swiss Institute of Bioinformatics and the EMBL outstation -   |        |     |
| CC                              | the European Bioinformatics Institute. There are no restrictions on its   |        |     |
| CC                              | use by non-profit institutions as long as its content is in no way  |        |     |
| CC                              | modified and this statement is not removed. Usage by and for commercial   |        |     |
| CC                              | entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/</a> |        |     |
| CC                              | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).                               |        |     |
| CC                              | EMBL; X14432; CAA32597.1; -.  |        |     |
| DR                              | PIR; S08488; A60501.  |        |     |
| DR                              | HSSP; P07204; LEGT.   |        |     |
| DR                              | MGI; 98736; Tbbd.   |        |     |
| DR                              | GO; GO:0009790; P:embryonic development; IMP.   |        |     |
| DR                              | GO; GO:0007565; P:pregnancy; IMP.   |        |     |
| DR                              | InterPro; IPR000152; ASX_hydroxyl_S.  |        |     |
| DR                              | InterPro; IPR001881; EGF_Ca.  |        |     |
| DR                              | InterPro; IPR006209; EGF-like.  |        |     |
| DR                              | InterPro; IPR001304; Lectin C.  |        |     |
| DR                              | InterPro; IPR001491; Thrombomodulin.  |        |     |
| DR                              | Pfam; PF00008; EGF_3.   |        |     |
| DR                              | Pfam; PF00059; Lectin_c; 1.   |        |     |
| DR                              | PRINTS; PR00907; THROMBOMODULN.   |        |     |
| DR                              | SMART; SM00034; CLEF; 1.  |        |     |
| DR                              | SMART; SM00179; EGF_Ca; 1.  |        |     |
| DR                              | PROSITE; PS00010; ASX_HYDROXYL; 2.  |        |     |
| DR                              | PROSITE; PS00041; C-TYPE LECTIN_2; 1.   |        |     |
| DR                              | PROSITE; PS00022; EGF_1; FALSE_NEG.   |        |     |
| DR                              | PROSITE; PS01186; EGF_2; 3.   |        |     |
| DR                              | PROSITE; PS00026; EGF_3; 3.   |        |     |
| DR                              | PROSITE; PS01187; EGF_Ca; 2.  |        |     |
| KW                              | Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;   |        |     |
| KW                              | Glycoprotein; Signal; EGF-like domain.  |        |     |
| FT                              | SIGNAL  | 1      | 16  |
| FT                              | CHAIN   | 17     | 577 |
| FT                              | DOMAIN  | 17     | 517 |
| FT                              | TRANSMEM  | 518    | 541 |
| FT                              | DOMAIN  | 542    | 577 |
| FT                              | DOMAIN  | 240    | 280 |
| FT                              | DOMAIN  | 283    | 323 |
| FT                              | DOMAIN  | 324    | 362 |
| FT                              | DOMAIN  | 364    | 404 |
| FT                              | DOMAIN  | 403    | 439 |
| FT                              | DOMAIN  | 440    | 480 |
| FT                              | DISULFID  | 244    | 255 |
| FT                              | DISULFID  | 251    | 264 |
| FT                              | DISULFID  | 266    | 279 |
| FT                              | DISULFID  | 287    | 295 |
| FT                              | DISULFID  | 291    | 307 |
| FT                              | DISULFID  | 309    | 322 |
| FT                              | DISULFID  | 328    | 339 |
| FT                              | DISULFID  | 335    | 348 |
| FT                              | DISULFID  | 350    | 361 |
| FT                              | DISULFID  | 368    | 377 |
| FT                              | DISULFID  | 373    | 387 |
| FT                              | DISULFID  | 389    | 403 |
| FT                              | DISULFID  | 407    | 416 |
| FT                              | DISULFID  | 412    | 424 |
| FT                              | DISULFID  | 426    | 438 |
| FT                              | DISULFID  | 444    | 454 |
| FT                              | DISULFID  | 449    | 463 |
| FT                              | DISULFID  | 465    | 479 |
| FT                              | CARBOHYD  | 113    | 113 |
| FT                              | CARBOHYD  | 243    | 243 |
| FT                              | CARBOHYD  | 256    | 256 |



|   |  |   |           |   |              |
|---|--|---|-----------|---|--------------|
| FT  | CARBOHYD   | 408   | 408       | N-LINKED (GLCNAC...)                          | (POTENTIAL). |
| FT  | CARBOHYD   | 494   | 494       | O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY). |              |
| SEQ   | SEQUENCE   | 577   | AA; 61867 | MW; B20E50B0FE745014                          | CRC64;       |
| Query Match   |  |   |           |   |              |
| Best Local Similarity 65.7%; Score 1858; DB 1; Length 577;      |  |   |           |   |              |
| Matches 332; Conservative 45; Mismatches 118; Indels 8; Gaps 4; |  |   |           |   |              |
| QY  | 1  | APAPQPGGSCVCHDFCALYPGPATFALNASQICDGLRGHLMTVRSVAADVISLLNGD   | 60        |   |              |
| DB  | 19   | ALAKLQPTGSCVCHDFCALYPGPATFALNASQICDGLRGHLMTVRSVAADVISLLSQS  | 78        |   |              |
| QY  | 61   | G-GVGRRLMIGLQPLPGCGDKPLGRGOWTGTGNTSYSHWRLDNLGAPLCPLC        | 119       |   |              |
| DB  | 79   | SMDLGP---WIGLQPLQCGDDPVHLGRGOWTGTGNTSYSHWRLDNLGAPLCPLC      | 135       |   |              |
| QY  | 120  | VAVSAAEATVSEPIWEEQCEKADGFLCEFHFPATCRPLAVEP-GAAAAAASVITYGTP  | 178       |   |              |
| DB  | 136  | VTVSTATEAARCEPAWEKPCETETQFLCEFYFTASCRPLTVNTRDPEAAHISSTYNTP  | 195       |   |              |
| QY  | 179  | FAARGADPQALPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCSHAC  | 238       |   |              |
| DB  | 196  | FGVSGADFOTLPVGSSAAVPLGLQMLCTAPPAGVQGHWAREAPGAWDCSVENGCSHAC  | 255       |   |              |
| QY  | 239  | NAIPGAPRCQCPAGNALQADGSCSTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRL | 298       |   |              |
| DB  | 256  | NRSTNEPRCLCPDMQLQADGSCSTASATQSCNDLCEHFCVNPDPQSGSYSCMCTGYRL  | 315       |   |              |
| QY  | 299  | AADHRCEDVDCCILEPSPQPCQPCVNTGGGECPCYNYDLVDGECVPEVPDPFRANCEY  | 358       |   |              |
| DB  | 316  | AADHRCEDVDCCILEPSPQPCQPCVNTGGGECPCYNYDLVDGECVPEVPDPFRANCEY  | 375       |   |              |
| QY  | 359  | QCQPLNQTSLYLCVCAAGPAPIPHEPHRCQPCNCTACPADCDPNTQASCRCPGVILDDG | 418       |   |              |
| DB  | 376  | QCQVSPDYRICACAPGAPKAPKPEPHKCEMFCNETSCPADCDPNTQASCRCPGVILDDG | 435       |   |              |
| QY  | 419  | FICTDIDCEGFGGFCGCHNLPGTFECICGPDALRHIGTDCDSGKV---DGDGSGG     | 475       |   |              |
| DB  | 436  | SVCTDIDCEGFGGFCGCHNLPGTFECICGPDALRHIGTDCDSGKV---DGDGSGG     | 495       |   |              |
| QY  | 476  | EPSPSTPGSTLTPPAVGLVHSG 498                                  |           |   |              |
| DB  | 496  | EPSPSTPGSTLTPPAVGLVHSG 518                                  |           |   |              |
| RESULT 3  |  |   |           |   |              |
| TRMB  | BOVIN  | STANDARD;   | PRT;      | 356   | AA.          |
| ID  | TRMB   | BOVIN   |           |   |              |
| AC  | P06579;  |   |           |   |              |
| DT  | 01-JAN-1988  | (Rel. 06, Created)  |           |   |              |
| DT  | 01-JAN-1988  | (Rel. 06, Last sequence update)                             |           |   |              |
| DT  | 16-OCT-2001  | (Rel. 40, Last annotation update)                           |           |   |              |
| DE  | Thrombomodulin (fetomodulin) (TM) (Fragment).                        |   |           |   |              |
| GN  | THBD.  |   |           |   |              |
| OS  | Bos taurus (Bovine).   |   |           |   |              |
| OC  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |   |           |   |              |
| OC  | Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;    |   |           |   |              |
| OC  | Bovidae; Bovinae; Bos.   |   |           |   |              |
| OK  | NCBI_TaxID=9913;   |   |           |   |              |
| RN  | [1]  |   |           |   |              |
| RP  | SEQUENCE FROM N.A.   |   |           |   |              |
| RX  | MEDLINE=67067408; PubMed=3024152;                                    |   |           |   |              |
| RA  | Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;            |   |           |   |              |
| RT  | "Characterization of a thrombomodulin cDNA reveals structural        |   |           |   |              |
| RT  | similarity to the low density lipoprotein receptor."                 |   |           |   |              |
| RL  | Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838 (1986).                   |   |           |   |              |
| CC  | -!- FUNCTION: Thrombomodulin is a specific endothelial cell receptor |   |           |   |              |
| CC  | that forms a 1:1 stoichiometric complex with thrombin. This          |   |           |   |              |
| CC  | complex is responsible for the conversion of protein C to the        |   |           |   |              |
| CC  | activated protein C (protein Ca). Once evolved, protein Ca           |   |           |   |              |
| CC  | scissions the activated cofactors of the coagulation mechanism,      |   |           |   |              |
| CC  | factor Va and factor VIIIa, and thereby reduces the amount of        |   |           |   |              |

|   |  |   |     |
|---|--|---|-----|
| CC  | thrombin generated.  |   |     |
| CC  | -!- SUBCELLULAR LOCATION: Type I membrane protein.   |   |     |
| CC  | -!- TISSUE SPECIFICITY: Endothelial cells are unique in synthesizing thrombomodulin.   |   |     |
| CC  | -!- SIMILARITY: Contains 6 EGF-like domains.   |   |     |
| CC  | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ). |   |     |
| CC  | EMBL; M14657; AAA30785.1; -.   |   |     |
| DR  | PIR; A25918; A25918.   |   |     |
| DR  | HSP; P07204; 1TMR.   |   |     |
| DR  | InterPro; IPR000152; Asx hydroxyl_s.   |   |     |
| DR  | InterPro; IPR001881; EGF_Ca.   |   |     |
| DR  | InterPro; IPR006209; EGF-like.   |   |     |
| DR  | InterPro; IPR001491; Thrombomodulin.   |   |     |
| DR  | Pfam; PF00008; EGF; 4.   |   |     |
| DR  | PRINTS; PR00907; THROMBOMODULIN.   |   |     |
| DR  | SMART; SM00179; EGF CA; 1.   |   |     |
| DR  | PROSITE; PS00010; ASX_HYDROXYL; 2.   |   |     |
| DR  | PROSITE; PS00022; EGF_1; FALSE_NEG.  |   |     |
| DR  | PROSITE; PS01186; EGF_2; 3.  |   |     |
| DR  | PROSITE; PS00026; EGF_3; 3.  |   |     |
| DR  | PROSITE; PS01187; EGF CA; 2.   |   |     |
| DR  | Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane; Glycoprotein; EGF-like domain.   |   |     |
| KW  | NON TER  |   |     |
| FT  | DOMAIN 1 296 EXTRACELLULAR (POTENTIAL).  |   |     |
| FT  | TRANSMEM 297 320 POTENTIAL.  |   |     |
| FT  | DOMAIN 321 356 CYTOPLASMIC (POTENTIAL).  |   |     |
| FT  | DOMAIN 17 57 EGF-LIKE 1.   |   |     |
| FT  | DOMAIN 60 98 EGF-LIKE 2.   |   |     |
| FT  | DOMAIN 99 137 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).   |   |     |
| FT  | DOMAIN 139 179 EGF-LIKE 4.   |   |     |
| FT  | DOMAIN 178 214 EGF-LIKE 5.   |   |     |
| FT  | DOMAIN 215 254 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  |   |     |
| FT  | DISULFID 21 32 BY SIMILARITY.  |   |     |
| FT  | DISULFID 28 41 BY SIMILARITY.  |   |     |
| FT  | DISULFID 43 56 BY SIMILARITY.  |   |     |
| FT  | DISULFID 64 72 BY SIMILARITY.  |   |     |
| FT  | DISULFID 68 82 BY SIMILARITY.  |   |     |
| FT  | DISULFID 84 97 BY SIMILARITY.  |   |     |
| FT  | DISULFID 103 114 BY SIMILARITY.  |   |     |
| FT  | DISULFID 110 123 BY SIMILARITY.  |   |     |
| FT  | DISULFID 125 136 BY SIMILARITY.  |   |     |
| FT  | DISULFID 143 152 BY SIMILARITY.  |   |     |
| FT  | DISULFID 148 162 BY SIMILARITY.  |   |     |
| FT  | DISULFID 164 178 BY SIMILARITY.  |   |     |
| FT  | DISULFID 182 191 BY SIMILARITY.  |   |     |
| FT  | DISULFID 187 199 BY SIMILARITY.  |   |     |
| FT  | DISULFID 201 213 BY SIMILARITY.  |   |     |
| FT  | DISULFID 219 228 BY SIMILARITY.  |   |     |
| FT  | DISULFID 224 237 BY SIMILARITY.  |   |     |
| FT  | DISULFID 239 253 BY SIMILARITY.  |   |     |
| FT  | CARBOHYD 271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY SIMILARITY).   |   |     |
| SEQ   | SEQUENCE 356 AA; 37795 MW; 29B41F097ABE1093 CRC64;   |   |     |
| Query Match   |  | 38.7%; Score 1095.5; DB 1; Length 356;                      |     |
| Best Local Similarity   |  | 63.4%; Pred. No. 8e-66;                                     |     |
| Matches 189; Conservative 31; Mismatches 63; Indels 15; Gaps 6; |  |   |     |
| QY  | 211  | GAVQGHWAREAPGAWDCSVENGCSHACNAIPGAPRCQPCPAGALQADGSCSTASATQSC | 270 |
| Db  | 5  | GTEGHSREAPGAWACGVGGCQCKGAGAGNCLCPADALQADGSCSTASATQSC        | 64  |
| QY  | 271  | NDLCHEFCVNPDPQGSYSCMCTGYRLAADHRCEDVDCCILEPSPQPCQPCVNTQGGFE  | 330 |

"The DNA sequence and comparative analysis of human chromosome 20."; *Nature* 414:865-871 (2001).  
[4]  
SEQUENCE FROM N.A.  
TISSUE=Leukocyte;  
MEDLINE=22388257; PubMed=12477932;  
STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
"Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences."; *Proc. Natl. Acad. Sci. U.S.A.* 99:16899-16903 (2002).

## RESULT 4

[5] CHARACTERIZATION. PubMed=11994479;  
RX MEDLINE=21990337; PubMed=11994479;  
RX McGreal E.P., Ikewai N., Akateu H., Morgan B.P., Gasque P.;  
RT "Human C1qRp is identical with CD93 and the mNI-11 antigen but does  
RT not bind C1q.";  
RT J. Immunol. 168:5222-5232(2002).  
[6]  
RX O-GLYCOSYLATION.  
RX MEDLINE=99132777; PubMed=10092817;  
RX Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;  
RA "C1qRp is a heavily O-glycosylated cell surface protein involved in  
RA the regulation of phagocytic activity.";  
RT J. Immunol. 162:3583-3589(1999).  
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for  
CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant  
CC protein A (SPA). May mediate the enhancement of phagocytosis in  
CC monocytes and macrophages upon interaction with soluble defense  
CC collagens. May play a role in intercellular adhesion.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,  
CC platelets, cells of myeloid origin, such as monocytes and  
CC neutrophils. Not expressed in cells of lymphoid origin.  
CC -!- PM: N- and O-glycosylated.  
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.  
CC -!- SIMILARITY: Contains 5 EGF-like domains.  
CC -!- CAUTION: Has been sometimes referred to as a collectin receptor.  
CC -!- CAUTION: According to Ref.5, C1q is not a ligand for C1QR1.  
CC -!- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001);  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456.g.htm".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U94333; AAB53110.1; -;  
DR EMBL; AL118508; CAC00597.1; -;  
DR EMBL; BC028075; AAC08075.1; -;  
DR HSP; P35555; 1EMN.  
DR Genew; HGNC:15855; C1QR1.  
DR MIM; 120577; -;  
DR GO; GO:0016021; C:integral to membrane; IC.  
DR GO; GO:0004872; F:receptor activity; NAS.





OX NCBI\_TaxID=101116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PVG; TISSUE=Natural killer cells;  
 RX MEDLINE=20545218; PubMed=11093152;  
 RA Lovik G., Vaage J.T., Dissen E., Spierer C., Ryan J.C., Rolstad B.;  
 RT "Characterization and molecular cloning of rat ClqR, a receptor on NK  
 cells.";   
 RL Eur. J. Immunol. 30:3355-3362(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Lung;  
 RX MEDLINE=20507883; PubMed=10934210;  
 RA Dean Y.D., McGreal E.P., Akatsue H., Gasque P.;  
 RT "Molecular and cellular properties of the rat A4 antigen, a C-type  
 lectin-like receptor with structural homology to thrombomodulin.";   
 RL J. Biol. Chem. 275:34382-34392(2000).  
 CC -I- FUNCTION: Receptor (or element of a larger receptor complex) for  
 CC Clq, mannose-binding lectin (MBL2) and pulmonary surfactant  
 CC protein A (SPA). May mediate the enhancement of phagocytosis in  
 CC monocytes and macrophages upon interaction with soluble defense  
 CC collagens. May play a role in intercellular adhesion.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and  
 CC heart. Expressed at lower level in brain, thymus, liver, spleen,  
 CC intestine, kidney, adrenal gland, muscle and testis. Expressed on  
 CC endothelial cells, platelets, undifferentiated monocytes and  
 CC circulating natural killer cells.  
 CC -I- PTM: N- and O-glycosylated (by similarity).  
 CC -I- SIMILARITY: Contains 1 C-type lectin family domain.  
 CC -I- SIMILARITY: Contains 5 EGF-like domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF136537; AAC01572.1; -;  
 CC HGSP; AF160978; AAF80402.1; -;  
 CC HSP; P35555; 1EMN.  
 CC GO: GO:0016021; C: integral to membrane; ISS.  
 CC GO: GO:0004872; F: receptor activity; ISS.  
 CC GO: GO:0016337; P: cell-cell adhesion; ISS.  
 CC GO: GO:0042116; P: macrophage activation; ISS.  
 CC GO: GO:0006909; P: phagocytosis; ISS.  
 CC InterPro: IPR001881; EGF\_Ca.  
 CC InterPro: IPR006209; EGF\_like.  
 CC InterPro: IPR001304; Lectin\_C.  
 CC Pfam: PF00008; EGF; 4.  
 CC Pfam: PF00059; lectin\_c; 1.  
 CC SMART: SM00034; CLECT; 1.  
 CC SMART: SM00179; EGF\_CA; 3.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 3.  
 CC PROSITE: PS00615; C\_TYPE\_LECTIN\_1; FALSE\_NEG.  
 CC PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
 CC PROSITE: PS01186; EGF\_2; 3.  
 CC PROSITE: PS50026; EGF\_3; 4.  
 CC PROSITE: PS01187; EGF\_CA; 3.  
 CC Cell adhesion; Receptor; Repeat; Signal; Transmembrane;  
 CC EGF-like domain; Lectin; Glycoprotein.  
 KW SIGNAL  
 FT 1 23  
 FT CHAIN 24 643  
 FT DOMAIN 24 571  
 FT TRANSMEM 572 592  
 FT DOMAIN 593 643  
 FT DOMAIN 31 173  
 FT DOMAIN 257 298  
 FT DOMAIN 299 341  
 FT DOMAIN 342 381  
 FT DOMAIN 381 485  
 FT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 424 462 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DISULFID 261 272 BY SIMILARITY.  
 FT DISULFID 268 282 BY SIMILARITY.  
 FT DISULFID 284 297 BY SIMILARITY.  
 FT DISULFID 303 314 BY SIMILARITY.  
 FT DISULFID 308 325 BY SIMILARITY.  
 FT DISULFID 327 340 BY SIMILARITY.  
 FT DISULFID 346 355 BY SIMILARITY.  
 FT DISULFID 351 364 BY SIMILARITY.  
 FT DISULFID 366 380 BY SIMILARITY.  
 FT DISULFID 386 397 BY SIMILARITY.  
 FT DISULFID 393 406 BY SIMILARITY.  
 FT DISULFID 408 422 BY SIMILARITY.  
 FT DISULFID 428 437 BY SIMILARITY.  
 FT DISULFID 433 446 BY SIMILARITY.  
 FT DISULFID 448 461 BY SIMILARITY.  
 FT CARBOHYD 322 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 417 417 E -> K (IN REF. 2).  
 SQ SEQUENCE 643 AA; 68781 MW; 9AE4C933AD943DB6 CRC64;  
 Query Match 18.6%; Score 525.5; DB 1; Length 643;  
 Best Local Similarity 30.1%; Pred. No. 7.8e-28;  
 Matches 160; Conservative 53; Mismatches 196; Indels 123; Gaps 27;  
 QY 12 CVEHDFCALYPGATFLNASQICDGLRGLHMTVRSSVAA-----DVISLLIN-----GDGGV 63  
 DB 30 CEGTACTYAHWGKLSAAEAQHRCNENGNLAVYKSEEAHVQEAQLKTKAPSETKI 89  
 QY 64 GRRLWIGLQLPPGCGDKPKRLGPFQWVTGNNTSYSRWARLDLNG--APUCGPLCVA 121  
 DB 90 G--KFWIGLQREKQKCTYHDL-PMKGFVWGGEDITTSNWKASKSSCIKRCVSLILD 146  
 QY 122 VSAEATVSE--PIWEEQCEV-----KADGLCEHFHPATCRPLAY-EPGAAAVSI 173  
 DB 147 LSLKPH--PSHLPMHESPCGTPDAPGNSIEGFLKFNFKMGCSPLALGGFG-----QL 198  
 QY 174 TYGTPTAARGADFOALPVGSSAAVAPLGLQ-----LMCTAPPGAVQGHWAREAP--- 222  
 DB 199 TYTTPQATTSLSLKAVFFASVANV--CGDEAKTNYLYCKETAGV-FHWSSGGLCVS 256  
 QY 223 GAWDCSVENGCGCEHAC-NAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 281  
 DB 257 PKFGCSFNNGGCGQDCFEAGDGSFGCRPGPRLDLDTVCAS-----RNPCCSN 306  
 QY 282 PDQPG-----SYSCMCETGYRLAADOHCEVDVDCILLEPSPCPQPCVNTQGFEC 331  
 DB 307 FCTGGGCHSVPLSENYTCHCPRGYQLDSSQVHCVIDDEC--EDSPCDQECINTPGGFHC 364  
 QY 332 HCYPNVDLVG--ECVEPVDPCFRANCEYQCQPLNOTSILCVCAEGFAPIPHEPHRCQMF 389  
 DB 365 ECWVGQSSGSGKEACEDVDEC-----TAAYSPCAQG----- 396  
 QY 390 CNOTACPADCDNPQAS--CECPGKYL--DDGFICTIDIDECENGFCSCVCHNLPTFE 445  
 DB 397 CT-----NTDGSFYCSCKEYIMSGEDTQCEDIDEC--LGNPCDTLCINTDGSFR 445  
 QY 446 CIGPDSALARHIGTDCDSG-----KVDGDSGSGEPPTPGS 485  
 DB 446 CGCPAGFELAPN-GVSCRTGSMFSELPARPQKEDKDGKSTVPLTEMPFGS 496  
 RESULT 7  
 FB2\_MOUSE  
 ID FB2\_MOUSE STANDARD; PRT; 2907 AA.  
 AC Q61555; Q63957;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Fibrillin 2 precursor.  
 GN FB2 OR FB2-2.  
 OS Mus musculus (Mouse).

|    |  |    |          |      |      |                               |
|----|--|----|----------|------|------|-------------------------------|
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  | FT | DOMAIN   | 1066 | 1107 | EGF-LIKE 15, CALCIUM-BINDING. |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.   | FT | DOMAIN   | 1108 | 1150 | EGF-LIKE 16, CALCIUM-BINDING. |
| OX | NCBI TaxID=10090;  | FT | DOMAIN   | 1151 | 1192 | EGF-LIKE 17, CALCIUM-BINDING. |
| RP | SEQUENCE FROM N.A.   | FT | DOMAIN   | 1193 | 1234 | EGF-LIKE 18, CALCIUM-BINDING. |
| RX | MEDLINE=95263670; PubMed=7744963;  | FT | DOMAIN   | 1235 | 1275 | EGF-LIKE 19, CALCIUM-BINDING. |
| RA | Zhang H., Hu W., Ramirez F.;   | FT | DOMAIN   | 1276 | 1317 | EGF-LIKE 20, CALCIUM-BINDING. |
| RA | "developmental expression of fibrillin genes suggests heterogeneity  | FT | DOMAIN   | 1318 | 1359 | EGF-LIKE 21, CALCIUM-BINDING. |
| RT | of extracellular microfibrils.";   | FT | DOMAIN   | 1360 | 1400 | EGF-LIKE 22, CALCIUM-BINDING. |
| RL | J. Cell Biol. 129:1165-1176(1995).   | FT | DOMAIN   | 1401 | 1441 | EGF-LIKE 23, CALCIUM-BINDING. |
| RN | [2]  | FT | DOMAIN   | 1442 | 1483 | EGF-LIKE 24, CALCIUM-BINDING. |
| RP | SEQUENCE OF 210-317 FROM N.A.  | FT | DOMAIN   | 1484 | 1524 | EGF-LIKE 25, CALCIUM-BINDING. |
| RX | MEDLINE=94140368; PubMed=8307578;  | FT | REPEAT   | 1525 | 1565 | EGF-LIKE 26, CALCIUM-BINDING. |
| RA | Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,   | FT | DOMAIN   | 1566 | 1642 | TGFBP 4.                      |
| RA | Francis U.;  | FT | DOMAIN   | 1643 | 1684 | EGF-LIKE 27, CALCIUM-BINDING. |
| RT | "Fibrillin genes map to regions of conserved mouse/human syntenic on   | FT | REPEAT   | 1685 | 1726 | EGF-LIKE 28, CALCIUM-BINDING. |
| RT | mouse chromosomes 2 and 18.";  | FT | DOMAIN   | 1727 | 1800 | TGFBP 5.                      |
| RL | Genomics 18:667-672(1993).   | FT | DOMAIN   | 1801 | 1842 | EGF-LIKE 29, CALCIUM-BINDING. |
| CC | -I- FUNCTION: Structural component of connective tissue microfibrils   | FT | DOMAIN   | 1843 | 1884 | EGF-LIKE 30, CALCIUM-BINDING. |
| CC | that binds calcium. Fibrillin-2-containing microfibrils regulate   | FT | DOMAIN   | 1885 | 1926 | EGF-LIKE 31, CALCIUM-BINDING. |
| CC | the early process of elastic fiber assembly.   | FT | DOMAIN   | 1927 | 1965 | EGF-LIKE 32, CALCIUM-BINDING. |
| CC | -I- SIMILARITY: Contains 47 EGF-like domains.  | FT | DOMAIN   | 1966 | 2008 | EGF-LIKE 33, CALCIUM-BINDING. |
| CC | -I- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.   | FT | DOMAIN   | 2009 | 2048 | EGF-LIKE 34, CALCIUM-BINDING. |
| CC | -----  | FT | DOMAIN   | 2049 | 2090 | EGF-LIKE 35, CALCIUM-BINDING. |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration   | FT | REPEAT   | 2091 | 2163 | TGFBP 6.                      |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -  | FT | DOMAIN   | 2164 | 2205 | EGF-LIKE 36, CALCIUM-BINDING. |
| CC | the European Bioinformatics Institute. There are no restrictions on its  | FT | DOMAIN   | 2206 | 2245 | EGF-LIKE 37, CALCIUM-BINDING. |
| CC | use by non-profit institutions as long as its content is in no way   | FT | DOMAIN   | 2246 | 2286 | EGF-LIKE 38, CALCIUM-BINDING. |
| CC | modified and this statement is not removed. Usage by and for commercial  | FT | DOMAIN   | 2287 | 2330 | EGF-LIKE 39, CALCIUM-BINDING. |
| CC | entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> | FT | DOMAIN   | 2331 | 2372 | EGF-LIKE 40, CALCIUM-BINDING. |
| CC | or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).  | FT | REPEAT   | 2373 | 2441 | TGFBP 7.                      |
| CC | -----  | FT | DOMAIN   | 2442 | 2483 | EGF-LIKE 41, CALCIUM-BINDING. |
| DR | EMBL; L39790; AAA74908.1; -;   | FT | DOMAIN   | 2484 | 2524 | EGF-LIKE 42, CALCIUM-BINDING. |
| DR | EMBL; S63359; AAC60685.1; -;   | FT | DOMAIN   | 2525 | 2563 | EGF-LIKE 43, CALCIUM-BINDING. |
| DR | PIR; A57278; A57278.   | FT | DOMAIN   | 2564 | 2606 | EGF-LIKE 44, CALCIUM-BINDING. |
| DR | HSP; P35555; 1EMN.   | FT | DOMAIN   | 2607 | 2645 | EGF-LIKE 45, CALCIUM-BINDING. |
| DR | MGD; G0195490; Fbn2.   | FT | DOMAIN   | 2646 | 2687 | EGF-LIKE 46, CALCIUM-BINDING. |
| DR | GO; GO:0030326; P:limb morphogenesis; IMP.   | FT | DOMAIN   | 2688 | 2727 | EGF-LIKE 47, CALCIUM-BINDING. |
| DR | InterPro; IPR00152; Asx hydroxylase.   | FT | DISULFID | 115  | 124  | BY SIMILARITY.                |
| DR | InterPro; IPR001881; EGF_Ca.   | FT | DISULFID | 125  | 130  | BY SIMILARITY.                |
| DR | InterPro; IPR001438; EGF-II.   | FT | DISULFID | 131  | 141  | BY SIMILARITY.                |
| DR | InterPro; IPR006209; EGF-like.   | FT | DISULFID | 142  | 159  | BY SIMILARITY.                |
| DR | InterPro; IPR002212; Fibril-associ.  | FT | DISULFID | 160  | 175  | BY SIMILARITY.                |
| DR | Pfam; PF00008; EGF; 4.   | FT | DISULFID | 176  | 190  | BY SIMILARITY.                |
| DR | Pfam; PF00683; TB; 9.  | FT | DISULFID | 181  | 196  | BY SIMILARITY.                |
| DR | PRINTS; PR00101; EGF-BLOOD.  | FT | DISULFID | 197  | 207  | BY SIMILARITY.                |
| DR | SMART; SMO0179; EGF_CA; 43.  | FT | DISULFID | 208  | 292  | BY SIMILARITY.                |
| DR | PROSITE; PS00010; ASX HYDROXYL; 43.  | FT | DISULFID | 287  | 301  | BY SIMILARITY.                |
| DR | PROSITE; PS00022; EGF_1; 2.  | FT | DISULFID | 302  | 316  | BY SIMILARITY.                |
| DR | PROSITE; PS01186; EGF_2; 36.   | FT | DISULFID | 317  | 334  | BY SIMILARITY.                |
| DR | PROSITE; PS50026; EGF_3; 45.   | FT | DISULFID | 322  | 343  | BY SIMILARITY.                |
| DR | PROSITE; PS01187; EGF_CA; 43.  | FT | DISULFID | 329  | 343  | BY SIMILARITY.                |
| KW | Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  | FT | DISULFID | 345  | 358  | BY SIMILARITY.                |
| KW | Repeat; Signal; Multigene family.  | FT | DISULFID | 359  | 384  | BY SIMILARITY.                |
| FT | SIGNAL 1   | FT | DISULFID | 491  | 503  | BY SIMILARITY.                |
| FT | CHAIN 29   | FT | DISULFID | 498  | 512  | BY SIMILARITY.                |
| FT | DOMAIN 111   | FT | DISULFID | 513  | 526  | BY SIMILARITY.                |
| FT | DOMAIN 145   | FT | DISULFID | 527  | 542  | BY SIMILARITY.                |
| FT | DOMAIN 176   | FT | DISULFID | 532  | 551  | BY SIMILARITY.                |
| FT | DOMAIN 276   | FT | DISULFID | 537  | 551  | BY SIMILARITY.                |
| FT | DOMAIN 318   | FT | DISULFID | 553  | 566  | BY SIMILARITY.                |
| FT | REPEAT 360   | FT | DISULFID | 567  | 584  | BY SIMILARITY.                |
| FT | DOMAIN 487   | FT | DISULFID | 572  | 593  | BY SIMILARITY.                |
| FT | DOMAIN 528   | FT | DISULFID | 579  | 593  | BY SIMILARITY.                |
| FT | DOMAIN 568   | FT | DISULFID | 595  | 608  | BY SIMILARITY.                |
| FT | DOMAIN 610   | FT | DISULFID | 614  | 625  | BY SIMILARITY.                |
| FT | DOMAIN 651   | FT | DISULFID | 620  | 634  | BY SIMILARITY.                |
| FT | REPEAT 692   | FT | DISULFID | 635  | 649  | BY SIMILARITY.                |
| FT | DOMAIN 761   | FT | DISULFID | 649  | 666  | BY SIMILARITY.                |
| FT | DOMAIN 803   | FT | DISULFID | 655  | 666  | BY SIMILARITY.                |
| FT | DOMAIN 845   | FT | DISULFID | 661  | 675  | BY SIMILARITY.                |
| FT | DOMAIN 883   | FT | DISULFID | 677  | 690  | BY SIMILARITY.                |
| FT | DOMAIN 948   | FT | DISULFID | 765  | 777  | BY SIMILARITY.                |
| FT | REPEAT 990   | FT | DISULFID | 772  | 786  | BY SIMILARITY.                |
| FT | DOMAIN 1065  | FT | DISULFID | 788  | 801  | BY SIMILARITY.                |
| FT | DOMAIN 803   | FT | DISULFID | 807  | 819  | BY SIMILARITY.                |
| FT | DOMAIN 845   | FT | DISULFID | 828  | 843  | BY SIMILARITY.                |
| FT | DOMAIN 883   | FT | DISULFID | 830  |      |                               |





Query Match 13.0%; Score 368; DB 1; Length 2871;



Db 1310 GYSGKGGKGTCTD-INCEIGAHNCBHVACTWAGSKCSPGHWIGDKICTDLDECS 1368  
QY 381 HEPHRCQMFQNTACPADCPNTOAS--CECPGYILDGFICTDIDEC-EN----- 429  
Db 1369 NGTHMCSQH-----ADC-KNTMGSVRLCKEGY-TGDGFTCTDLDECSNLCGNGQ 1419  
QY 430 -----CGF-----CS-----GVCHNLGCTFECICGPPSALA 455  
Db 1420 CLNAPGGYRCEDMGFVPSADGKACEDIDCSLPNICVFTCHNLGFCBCEIGYELD 1479  
QY 456 RHIG-----TDCDSGK-VDGDSGSGBPSPPTGSLTTPPAVLV 495  
Db 1480 RSGNCTDVNECLDPTTCISGNCVNTPGSYTCDPPD-----FELNPRVGCV 1527  
RESULT 9  
FBL2 HUMAN STANDARD; PRT; 1184 AA.  
AC P98055; AC 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Fubulin-2 precursor.  
GN FBLN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Fibroblast;  
RX MEDLINE=95104855; PubMed=7806230;  
RA Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R.,  
Chen M.-L.;  
RT "Fubulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping  
of the gene on human and mouse chromosomes.";  
RL Genomics 22:425-430(1994).  
[2]  
RN DEVELOPMENTAL STAGE.  
RP  
RX MEDLINE=96301678; PubMed=8737292;  
RA Miosge N., Gotz W., Sasaki T., Chu M.-L., Timpl R., Herken R.;  
RT "The extracellular matrix proteins fibulin-1 and fibulin-2 in the  
early human embryo.";  
RL Histochem. J. 28:109-116(1996).  
CC -!- FUNCTION: Its binding to fibronectin and some other ligands is  
calcium dependent.  
CC -!- SUBUNIT: Homotrimer; disulfide-linked. Interacts with LAMA2 (By  
similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
CC -!- TISSUE SPECIFICITY: Component of both basement membranes and other  
connective tissues. Expressed in heart, placenta and ovary.  
CC -!- DEVELOPMENTAL STAGE: Widely expressed during embryonic  
development. Primarily detected within the neuroepithelium, spinal  
ganglia and peripheral nerves.  
CC -!- SIMILARITY: Belongs to the fibulin family.  
CC -!- SIMILARITY: Contains 3 anaphylatoxin-like domains.  
CC -!- SIMILARITY: Contains 11 EGF-like domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>  
or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL; X82494; CAA57876.1; -  
DR PIR; A55184; A55184.  
DR HSSP; P00736; IAPQ.  
DR Genew; HGNC:3601; FBLN2.  
DR MIM; 135821; -  
GO; GO:0005578; C:extracellular matrix; TAS.

DR GO; GO:0005509; F:calcium ion binding; TAS.  
DR GO; GO:0005201; F:extracellular matrix structural constituent; TAS.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR000152; Asx hydroxyl S.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF-like.  
DR Pfam; PF01821; ANATO; 2.  
DR Pfam; PF00008; EGF; 5.  
DR SMART; SM00104; ANATO; 3.  
DR SMART; SM00179; EGF CA; 9.  
DR PROSITE; PS01177; ANAPHYLATOXIN 1; 3.  
DR PROSITE; PS01178; ANAPHYLATOXIN 2; 3.  
DR PROSITE; PS00010; ASX HYDROXYL; 5.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 5.  
DR PROSITE; PS00026; EGF\_3; 4.  
DR PROSITE; PS01187; EGF CA; 9.  
KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
KW Calcium-binding; Repeat.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 1184 FIBULIN-2.  
FT DOMAIN 28 444 N.  
FT DOMAIN 28 177 SUBDOMAIN NA (CYS-RICH).  
FT DOMAIN 178 444 SUBDOMAIN NB (CYS-FREE).  
FT DOMAIN 445 480 ANAPHYLATOXIN-LIKE 1.  
FT DOMAIN 488 519 ANAPHYLATOXIN-LIKE 2.  
FT DOMAIN 521 553 ANAPHYLATOXIN-LIKE 3.  
FT DOMAIN 604 645 EGF-LIKE 1, CALCIUM-BINDING.  
FT DOMAIN 679 718 EGF-LIKE 2.  
FT DOMAIN 719 763 EGF-LIKE 3, CALCIUM-BINDING.  
FT DOMAIN 764 809 EGF-LIKE 4, CALCIUM-BINDING.  
FT DOMAIN 810 857 EGF-LIKE 5, CALCIUM-BINDING.  
FT DOMAIN 858 900 EGF-LIKE 6, CALCIUM-BINDING.  
FT DOMAIN 901 942 EGF-LIKE 7, CALCIUM-BINDING.  
FT DOMAIN 943 981 EGF-LIKE 8, CALCIUM-BINDING.  
FT DOMAIN 982 1024 EGF-LIKE 9, CALCIUM-BINDING.  
FT DOMAIN 1025 1069 EGF-LIKE 10, CALCIUM-BINDING.  
FT DOMAIN 1070 1184 DOMAIN III.  
FT DISULFID 445 472 BY SIMILARITY.  
FT DISULFID 446 479 BY SIMILARITY.  
FT DISULFID 459 480 BY SIMILARITY.  
FT DISULFID 489 518 BY SIMILARITY.  
FT DISULFID 502 519 BY SIMILARITY.  
FT DISULFID 521 545 BY SIMILARITY.  
FT DISULFID 522 552 BY SIMILARITY.  
FT DISULFID 535 553 BY SIMILARITY.  
FT DISULFID 608 620 BY SIMILARITY.  
FT DISULFID 616 629 BY SIMILARITY.  
FT DISULFID 631 644 BY SIMILARITY.  
FT DISULFID 683 693 BY SIMILARITY.  
FT DISULFID 689 702 BY SIMILARITY.  
FT DISULFID 704 717 BY SIMILARITY.  
FT DISULFID 723 736 BY SIMILARITY.  
FT DISULFID 730 745 BY SIMILARITY.  
FT DISULFID 751 762 BY SIMILARITY.  
FT DISULFID 768 781 BY SIMILARITY.  
FT DISULFID 775 790 BY SIMILARITY.  
FT DISULFID 796 808 BY SIMILARITY.  
FT DISULFID 814 827 BY SIMILARITY.  
FT DISULFID 821 836 BY SIMILARITY.  
FT DISULFID 843 856 BY SIMILARITY.  
FT DISULFID 862 875 BY SIMILARITY.  
FT DISULFID 869 884 BY SIMILARITY.  
FT DISULFID 886 899 BY SIMILARITY.  
FT DISULFID 905 917 BY SIMILARITY.  
FT DISULFID 913 926 BY SIMILARITY.  
FT DISULFID 928 941 BY SIMILARITY.  
FT DISULFID 947 956 BY SIMILARITY.  
FT DISULFID 952 965 BY SIMILARITY.  
FT DISULFID 967 980 BY SIMILARITY.  
FT DISULFID 986 998 BY SIMILARITY.  
FT DISULFID 994 1007 BY SIMILARITY.  
FT DISULFID 1009 1023 BY SIMILARITY.

FT DISULFID 1029 1042 BY SIMILARITY.  
 FT DISULFID 1036 1051 BY SIMILARITY.  
 FT DISULFID 1056 1068 BY SIMILARITY.  
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F9ECSD CRC64;  
 Query Match 13.0%; Score 367.5; DB 1; Length 1184;  
 Best Local Similarity 30.0%; Pred. No. 4.1e-17;  
 Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
 QY 136 EEOQCEVKAD-GFLCEFHPRATC-----RPLAV-----EPGAAAAAASITYGTTPFAARG 183  
 Db 531 EGQSCSNPNLGPCH-HVMSLSCBEEPLIVEVRPPEPAAPRVS-----EAEH 582  
 QY 184 ADFOALPVGSSAAV---APLGLQLMCTAPPGVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 Db 583 AGREALSLGTEAELPNSLPDGDQDECLLPGL-----COHLCLIN 622  
 QY 241 IPGAPRCOCFAGALQADGRSC-----TASAT-----QS 269  
 Db 623 TVGSYHACAPGFSLODGGSTCRPEGHPPQPEAPQBPALKSEFSQVASTPIPLPQPNP 652  
 QY 270 CND--LCEHFCVNPDPQSGYSQWCTGYRLAADOHRCEVDVDCILEPSPCP--QRQCVNT 325  
 Db 683 CKDNGPKQVC--STVGGSAIASCFFGYAIMADGVSCEDINECVDTLHTCSRGEHCVT 739  
 QY 326 QGGEFECH---CYPNYDLVDEGCEVPEVDPFRANCEYQCQPLNQTSLYCV-----371  
 Db 740 LGSFHCYKALTCBFGYALKDGECE-EDVDEC--AMGHTCQF-----GFLCQNTKGSFYCQA 792  
 QY 372 ---CABGFAPIPH-----EPHRCQMFQNTACPADCDNTQASCCEPGY-I 414  
 Db 793 RQRCMDGFLQDPGNCVDINECTSLSEPCFPGSCINTVGSYTCORNPLI---CARGYHA 849  
 QY 415 LDGDFICTDIDECENGFCSCG---VCHNLPGTPECIC 448  
 Db 850 SDDGAKVDVNEGETGVHRCGEQVCHNLPGSYRCDC 886  
 RESULT 10  
 FBNI\_PIG  
 ID FBNI\_PIG STANDARD; PRT; 2871 AA.  
 AC QSTV36;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBNI.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=Lung;  
 RX MEDLINE=99156858; PubMed=10036187;  
 RA Biery N.J., Eldadah Z.A., Moore C.S., Stetten G., Spencer F.,  
 RA Dietz H.C.;  
 RT "Revised genomic organization of FBNI and significance for regulated  
 RT gene expression";  
 RL Genomics 56:70-77(1999).  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 CC that binds calcium. Fibrillin-1-containing microfibrils provide  
 CC long-term force bearing structural support.  
 CC -!- PFM: Forms intermolecular disulfide bonds either with other  
 CC fibrillin-1 molecules or with other components of the  
 CC microfibrils.  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF073800; RAD50328.1; --  
 DR HSP; P35555; IAPU.  
 DR InterPro: IPR000152; Asx\_hydroxyl\_S.  
 DR InterPro: IPR001881; EGF Ca.  
 DR InterPro: IPR006209; EGF like.  
 DR InterPro: IPR002212; Fibril-assoc.  
 DR Pfam: PF00008; EGF; 44.  
 DR Pfam: PF00883; TB; 9.  
 DR SMART; SM00179; EGF\_CA; 40.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 41.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 36.  
 DR PROSITE; PS00026; EGF\_3; 43.  
 DR PROSITE; PS01187; EGF\_CA; 41.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 2871 FIBRILLIN 1.  
 FT DOMAIN 81 112 EGF-LIKE 1.  
 FT DOMAIN 115 146 EGF-LIKE 2.  
 FT DOMAIN 147 178 EGF-LIKE 3.  
 FT DOMAIN 246 287 EGF-LIKE 4. CALCIUM-BINDING.  
 FT DOMAIN 288 329 EGF-LIKE 5. CALCIUM-BINDING.  
 FT REPEAT 330 390 TGFBP 1.  
 FT DOMAIN 392 446 PRO-RICH.  
 FT DOMAIN 449 489 EGF-LIKE 6.  
 FT DOMAIN 490 529 EGF-LIKE 7. CALCIUM-BINDING.  
 FT DOMAIN 530 571 EGF-LIKE 8. CALCIUM-BINDING.  
 FT DOMAIN 572 612 EGF-LIKE 9. CALCIUM-BINDING.  
 FT DOMAIN 613 653 EGF-LIKE 10. CALCIUM-BINDING.  
 FT REPEAT 654 722 TGFBP 2.  
 FT DOMAIN 723 764 EGF-LIKE 11. CALCIUM-BINDING.  
 FT DOMAIN 765 806 EGF-LIKE 12. CALCIUM-BINDING.  
 FT DOMAIN 807 846 EGF-LIKE 13. CALCIUM-BINDING.  
 FT DOMAIN 910 951 EGF-LIKE 14. CALCIUM-BINDING.  
 FT REPEAT 952 1027 TGFBP 3.  
 FT DOMAIN 1028 1069 EGF-LIKE 15. CALCIUM-BINDING.  
 FT DOMAIN 1070 1112 EGF-LIKE 16. CALCIUM-BINDING.  
 FT DOMAIN 1113 1154 EGF-LIKE 17. CALCIUM-BINDING.  
 FT DOMAIN 1155 1196 EGF-LIKE 18. CALCIUM-BINDING.  
 FT DOMAIN 1197 1237 EGF-LIKE 19. CALCIUM-BINDING.  
 FT DOMAIN 1238 1279 EGF-LIKE 20. CALCIUM-BINDING.  
 FT DOMAIN 1280 1321 EGF-LIKE 21. CALCIUM-BINDING.  
 FT DOMAIN 1322 1362 EGF-LIKE 22. CALCIUM-BINDING.  
 FT DOMAIN 1363 1403 EGF-LIKE 23. CALCIUM-BINDING.  
 FT DOMAIN 1404 1445 EGF-LIKE 24. CALCIUM-BINDING.  
 FT DOMAIN 1446 1486 EGF-LIKE 25. CALCIUM-BINDING.  
 FT DOMAIN 1487 1527 EGF-LIKE 26. CALCIUM-BINDING.  
 FT REPEAT 1528 1605 TGFBP 4.  
 FT DOMAIN 1606 1647 EGF-LIKE 27. CALCIUM-BINDING.  
 FT DOMAIN 1648 1688 EGF-LIKE 28. CALCIUM-BINDING.  
 FT REPEAT 1689 1765 TGFBP 5.  
 FT DOMAIN 1766 1807 EGF-LIKE 29. CALCIUM-BINDING.  
 FT DOMAIN 1808 1848 EGF-LIKE 30. CALCIUM-BINDING.  
 FT DOMAIN 1849 1890 EGF-LIKE 31. CALCIUM-BINDING.  
 FT DOMAIN 1891 1929 EGF-LIKE 32. CALCIUM-BINDING.  
 FT DOMAIN 1930 1972 EGF-LIKE 33. CALCIUM-BINDING.  
 FT DOMAIN 1973 2012 EGF-LIKE 34. CALCIUM-BINDING.  
 FT DOMAIN 2013 2054 EGF-LIKE 35. CALCIUM-BINDING.  
 FT REPEAT 2055 2126 TGFBP 6.  
 FT DOMAIN 2127 2165 EGF-LIKE 36. CALCIUM-BINDING.  
 FT DOMAIN 2166 2205 EGF-LIKE 37. CALCIUM-BINDING.  
 FT DOMAIN 2206 2246 EGF-LIKE 38. CALCIUM-BINDING.  
 FT DOMAIN 2247 2290 EGF-LIKE 39. CALCIUM-BINDING.  
 FT DOMAIN 2291 2332 EGF-LIKE 40. CALCIUM-BINDING.  
 FT REPEAT 2333 2401 TGFBP 7.

```
FT DOMAIN 2402 EGF-LIKE 41, CALCIUM-BINDING. 2443
FT DOMAIN 2444 EGF-LIKE 42, CALCIUM-BINDING. 2484
FT DOMAIN 2485 EGF-LIKE 43, CALCIUM-BINDING. 2523
FT DOMAIN 2524 EGF-LIKE 44, CALCIUM-BINDING. 2566
FT DOMAIN 2567 EGF-LIKE 45, CALCIUM-BINDING. 2606
FT DOMAIN 2607 EGF-LIKE 46, CALCIUM-BINDING. 2647
FT DOMAIN 2648 EGF-LIKE 47, CALCIUM-BINDING. 2687
FT DISULFID 85 94
FT DISULFID 89 100
FT DISULFID 102 111
FT DISULFID 119 129
FT DISULFID 123 134
FT DISULFID 136 145
FT DISULFID 150 160
FT DISULFID 154 166
FT DISULFID 168 177
FT DISULFID 250 262
FT DISULFID 257 271
FT DISULFID 273 286
FT DISULFID 292 304
FT DISULFID 299 313
FT DISULFID 315 328
FT DISULFID 453 465
FT DISULFID 460 474
FT DISULFID 476 488
FT DISULFID 494 504
FT DISULFID 499 513
FT DISULFID 515 528
FT DISULFID 534 546
FT DISULFID 541 555
FT DISULFID 557 570
FT DISULFID 576 587
FT DISULFID 582 596
FT DISULFID 598 611
FT DISULFID 617 628
FT DISULFID 623 637
FT DISULFID 639 652
FT DISULFID 727 739
FT DISULFID 734 748
FT DISULFID 750 763
FT DISULFID 769 781
FT DISULFID 776 790
FT DISULFID 792 805
FT DISULFID 811 821
FT DISULFID 816 830
FT DISULFID 832 845
FT DISULFID 914 926
FT DISULFID 921 935
FT DISULFID 937 950
FT DISULFID 1032 1044
FT DISULFID 1039 1053
FT DISULFID 1055 1068
FT DISULFID 1074 1086
FT DISULFID 1081 1095
FT DISULFID 1097 1111
FT DISULFID 1117 1129
FT DISULFID 1124 1138
FT DISULFID 1140 1153
FT DISULFID 1159 1171
FT DISULFID 1201 1212
FT DISULFID 1208 1221
FT DISULFID 1223 1236
FT DISULFID 1242 1254
FT DISULFID 1249 1263
FT DISULFID 1265 1278
FT DISULFID 1284 1296
FT DISULFID 1291 1305
FT DISULFID 1307 1320
FT DISULFID 1326 1339
FT DISULFID 1333 1348
FT DISULFID 1350 1361
FT DISULFID 1367 1380
FT DISULFID 1374 1389

Query Match 12.9%; Score 366; DB 1; Length 2871;
Best Local Similarity 30.6%; Pred. No. 1.2e-16;
Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

QY 226 DCSVENGGECHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCRHFCVNP-- 282
Db 1200 ECSIMMGCGCTCTNSEGSEYSCQCPGALPQPSCT-----DIDE--CEDNPNIC 1249
QY 283 -----DQPSYSCMCETGYRLAADQHRCEVDVDDCILEPSPC-PQRCVNTQGFECYCYP 335
Db 1250 DGGQCTNIPGEYRCLCYDGFMASEDMKTCVDVNECDLNPNCILSGTCENTKGSFICHCDM 1309
QY 336 NYDLVDGE--CVEPVDPC--FRANCEYQCOPLNOT--SYLCVCAEFA-----PIP 380
Db 1310 GYSGKKGKTCCTD--INECEIGAHCNDRHACVNTAGSNFSCSPGWIGDGIKTDLDECS 1368
QY 381 HEPHRCQMFQNTACPADCPNTOAS--CECPGYILDDGFICTDIDEC-EN----- 429
Db 1369 NGTHMCSQH-----ADC-KNTMGSVYRLCKEGY-TGDGFTCADLDECSNVKLCGNVQ 1419
QY 430 -----GGF-----CS-----CVCHNLPGTFECICGPDLSALA 455
Db 1420 CLYAPGGYHCEYDMGFVPSADRKSCVDSDECSLPNTICVFGTCHNLPGLFRCEC----- 1472
QY 456 RHGTDCDCKVDGSGSGSGEPSPS-----PTPGS-----TLTPPAVGLV 495
Db 1473 -EIGYELDRSGGNCCTDVNECLEPPTCISGNCVNTPGSYTCVCPDFELNTRVGCV 1527

RESULT 11
FBN2 HUMAN ID_FBN2 HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguinetti C.,
RA Bonadio J., Mecham R.P., Ramirez F.;
RA "Structure and expression of fibrillin-2, a novel microfibrillar
```

RT component preferentially located in elastic matrices.";  
 RL J. Cell Biol. 124:855-863 (1994).  
 RN [2]  
 RP SEQUENCE OF 752-1505 FROM N.A.  
 RX MEDLINE=91304567; PubMed=1852206;  
 RA Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,  
 RA Tsipouras P., Ramirez F., Hollister D.;  
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
 RT two different fibrillin genes.";  
 RL Nature 352:330-334 (1991).  
 RN [3]  
 RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.  
 RX MEDLINE=96083599; PubMed=7493032;  
 RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;  
 RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,  
 RT congenital contractural arachnodactyly.";  
 RL Nat. Genet. 11:456-458 (1995).  
 RN [4]  
 RP VARIANT CCA HIS-1114.  
 RX MEDLINE=98407789; PubMed=9737771;  
 RA Babcock D., Gaerner C., Francke U., Maalen C.;  
 RT "A single mutation that results in an app-to-his substitution and  
 RT partial exon skipping in a family with congenital contractual  
 RT arachnodactyly.";  
 RL Hum. Genet. 103:22-28 (1998).  
 RN [5]  
 RP VARIANTS CCA PHE-1141 AND TRP-1252.  
 RX MEDLINE=20259236; PubMed=10797416;  
 RA Belleh S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,  
 RA Godfrey M.;  
 RT "Two novel fibrillin-2 mutations in congenital contractual  
 RT arachnodactyly.";  
 RL Am. J. Med. Genet. 92:7-12 (2000).  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 CC that binds calcium. Fibrillin-2-containing microfibrils regulate  
 CC the early process of elastic fiber assembly.  
 CC -!- DISEASE: Defects in FBN2 are the cause of congenital contractual  
 CC arachnodactyly (CCA) [MIM:121050]; also known as Beals syndrome.  
 CC CCA is phenotypically similar to Marfan syndrome, but does not  
 CC effect the aorta and the eyes.  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- DATABASE: Contains 7 TGF-beta binding protein (TGFBP) domains.  
 CC -!- DATABASE: NAME=Elastic Fiber Homepage; NOTE=Fibrillin 2 page;  
 CC WWW="http://ef.wustl.edu/genes/FBN2.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U03272; AAA18950.1; -;  
 CC EMBL; X62009; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; A54105; A54105.  
 CC HSP; P35555; IEMN.  
 CC Genew; HGNC:3604; FBN2.  
 CC MIM; 121050; -;  
 CC GO; GO:0005578; C:extracellular matrix; TAS.  
 CC GO; GO:0005201; F:extracellular matrix structural constituent; TAS.  
 CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.  
 CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR001512; Asx hydroxyl\_s.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR001438; EGF\_I1.  
 CC InterPro; IPR006209; EGF\_Like.  
 CC InterPro; IPR002212; Fibril-assoc.  
 CC Pfam; PF00008; EGF; 45.  
 CC Pfam; PF00683; TB; 9.  
 CC PRINTS; PRO0010; EGFBL00D.  
 CC SMART; SM00179; EGF\_CA; 43.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 43.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 37.  
 DR PROSITE; PS00026; EGF\_3; 45.  
 DR PROSITE; PS01187; EGF\_CA; 42.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.  
 FT SIGNAL 1 28  
 FT CHAIN 29 2911 FIBRILLIN 2.  
 FT DOMAIN 111 142 EGF-LIKE 1.  
 FT DOMAIN 145 176 EGF-LIKE 2.  
 FT DOMAIN 176 207 EGF-LIKE 3.  
 FT DOMAIN 275 316 EGF-LIKE 4. CALCIUM-BINDING.  
 FT DOMAIN 317 358 EGF-LIKE 5. CALCIUM-BINDING.  
 FT REPEAT 359 425 TGFBP 1.  
 FT DOMAIN 493 533 EGF-LIKE 6.  
 FT DOMAIN 534 573 EGF-LIKE 7. CALCIUM-BINDING.  
 FT DOMAIN 574 615 EGF-LIKE 8. CALCIUM-BINDING.  
 FT DOMAIN 616 656 EGF-LIKE 9. CALCIUM-BINDING.  
 FT DOMAIN 657 697 EGF-LIKE 10. CALCIUM-BINDING.  
 FT REPEAT 766 766 TGFBP 2.  
 FT DOMAIN 767 808 EGF-LIKE 11. CALCIUM-BINDING.  
 FT DOMAIN 809 850 EGF-LIKE 12. CALCIUM-BINDING.  
 FT DOMAIN 851 890 EGF-LIKE 13. CALCIUM-BINDING.  
 FT DOMAIN 954 995 EGF-LIKE 14. CALCIUM-BINDING.  
 FT REPEAT 996 1071 TGFBP 3.  
 FT DOMAIN 1072 1113 EGF-LIKE 15. CALCIUM-BINDING.  
 FT DOMAIN 1114 1156 EGF-LIKE 16. CALCIUM-BINDING.  
 FT DOMAIN 1157 1198 EGF-LIKE 17. CALCIUM-BINDING.  
 FT DOMAIN 1199 1240 EGF-LIKE 18. CALCIUM-BINDING.  
 FT DOMAIN 1241 1281 EGF-LIKE 19. CALCIUM-BINDING.  
 FT DOMAIN 1282 1323 EGF-LIKE 20. CALCIUM-BINDING.  
 FT DOMAIN 1324 1365 EGF-LIKE 21. CALCIUM-BINDING.  
 FT DOMAIN 1366 1406 EGF-LIKE 22. CALCIUM-BINDING.  
 FT DOMAIN 1407 1447 EGF-LIKE 23. CALCIUM-BINDING.  
 FT DOMAIN 1448 1489 EGF-LIKE 24. CALCIUM-BINDING.  
 FT DOMAIN 1490 1530 EGF-LIKE 25. CALCIUM-BINDING.  
 FT DOMAIN 1531 1571 EGF-LIKE 26. CALCIUM-BINDING.  
 FT REPEAT 1572 1649 TGFBP 4.  
 FT DOMAIN 1649 1690 EGF-LIKE 27. CALCIUM-BINDING.  
 FT DOMAIN 1691 1732 EGF-LIKE 28. CALCIUM-BINDING.  
 FT REPEAT 1733 1806 TGFBP 5.  
 FT DOMAIN 1807 1848 EGF-LIKE 29. CALCIUM-BINDING.  
 FT DOMAIN 1849 1890 EGF-LIKE 30. CALCIUM-BINDING.  
 FT DOMAIN 1891 1932 EGF-LIKE 31. CALCIUM-BINDING.  
 FT DOMAIN 1933 1971 EGF-LIKE 32. CALCIUM-BINDING.  
 FT DOMAIN 1972 2014 EGF-LIKE 33. CALCIUM-BINDING.  
 FT DOMAIN 2015 2054 EGF-LIKE 34. CALCIUM-BINDING.  
 FT DOMAIN 2055 2096 EGF-LIKE 35. CALCIUM-BINDING.  
 FT REPEAT 2097 2169 TGFBP 6.  
 FT DOMAIN 2170 2211 EGF-LIKE 36. CALCIUM-BINDING.  
 FT DOMAIN 2212 2251 EGF-LIKE 37. CALCIUM-BINDING.  
 FT DOMAIN 2252 2292 EGF-LIKE 38. CALCIUM-BINDING.  
 FT DOMAIN 2293 2336 EGF-LIKE 39. CALCIUM-BINDING.  
 FT DOMAIN 2337 2378 EGF-LIKE 40. CALCIUM-BINDING.  
 FT REPEAT 2379 2447 TGFBP 7.  
 FT DOMAIN 2448 2489 EGF-LIKE 41. CALCIUM-BINDING.  
 FT DOMAIN 2490 2530 EGF-LIKE 42. CALCIUM-BINDING.  
 FT DOMAIN 2531 2569 EGF-LIKE 43. CALCIUM-BINDING.  
 FT DOMAIN 2570 2612 EGF-LIKE 44. CALCIUM-BINDING.  
 FT DOMAIN 2613 2652 EGF-LIKE 45. CALCIUM-BINDING.  
 FT DOMAIN 2653 2693 EGF-LIKE 46. CALCIUM-BINDING.  
 FT DOMAIN 2694 2733 EGF-LIKE 47. CALCIUM-BINDING.  
 FT DISULFID 115 124 BY SIMILARITY.  
 FT DISULFID 119 130 BY SIMILARITY.  
 FT DISULFID 132 141 BY SIMILARITY.  
 FT DISULFID 149 159 BY SIMILARITY.  
 FT DISULFID 153 164 BY SIMILARITY.  
 FT DISULFID 166 175 BY SIMILARITY.  
 FT DISULFID 180 190 BY SIMILARITY.  
 FT DISULFID 184 195 BY SIMILARITY.  
 FT DISULFID 197 206 BY SIMILARITY.  
 FT DISULFID 291 291 BY SIMILARITY.  
 FT DISULFID 286 286 BY SIMILARITY.

10-OCT-2003 (Rel. 42, Last annotation update)

DE Fibrillin 1 precursor.  
FN1 OR FBN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Fibrillin.  
RC MEDLINE=93372850; PubMed=8364578;  
RX Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,  
RA Pangillan T., Bonadio J.;  
RA "Genomic organization of the sequence coding for fibrillin, the  
RT defective gene product in Marfan syndrome.";  
RN Hum. Mol. Genet. 2:961-968(1993).  
[2]  
RN SEQUENCE OF 1-932 FROM N.A.  
RP TISSUE=Fibroblast, and Placenta;  
RC MEDLINE=94010947; PubMed=7691719;  
RX Corson G.M., Chalberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;  
RA "Fibrillin binds calcium and is coded by cDNAs that reveal a  
RT multidomain structure and alternatively spliced exons at the 5'  
RT end.";  
RN Genomics 17:476-484(1993).  
[3]  
RN SEQUENCE OF 899-2871 FROM N.A.  
RP MEDLINE=91304568; PubMed=1852207;  
RX Maslen C.L., Corson G.M., Maddox B.K., Ghanville R.W., Sakai L.Y.;  
RA "Partial sequence of a candidate gene for the Marfan syndrome.";  
RN Nature 352:334-337(1991).  
[4]  
RN SEQUENCE OF 813-1313 FROM N.A.  
RP MEDLINE=91304567; PubMed=1852206;  
RX Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,  
RA Espoules P., Ramirez F., Hollister D.W.;  
RA "Linkage of Marfan syndrome and a phenotypically related disorder to  
RT two different fibrillin genes.";  
RN Nature 352:330-334(1991).  
[5]  
RN CHARACTERIZATION.  
RP MEDLINE=91317849; PubMed=1860873;  
RX Sakai L.Y., Keene D.R., Ghanville R.W., Bachinger H.P.;  
RA "Purification and partial characterization of fibrillin, a cysteine-  
RT rich structural component of connective tissue microfibrils";  
RN J. Biol. Chem. 266:14763-14770(1991).  
[6]  
RN STRUCTURE BY NMR OF 2054-2125.  
RP MEDLINE=98031893; PubMed=9362480;  
RX Yuan X., Downing A.K., Knott V., Handford P.A.;  
RA "Solution structure of the transforming growth factor beta-binding  
RT protein-like module, a domain associated with matrix fibrils.";  
RN EMBO J. 16:6659-6666(1997).  
[7]  
RN STRUCTURE BY NMR OF 2124-2205.  
RP MEDLINE=96144829; PubMed=8568869;  
RX Knott V., Downing A.K., Cady C.M., Handford P.A.;  
RA "Calcium binding properties of an epidermal growth factor-like domain  
RT pair from human fibrillin-1";  
RN J. Mol. Biol. 255:22-27(1996).  
[8]  
RN STRUCTURE BY NMR OF 2124-2205.  
RP MEDLINE=96222301; PubMed=8653794;  
RX Downing A.K., Knott V., Werner J.M., Cady C.M., Campbell I.D.,  
RA Handford P.A.;  
RA "Solution structure of a pair of calcium-binding epidermal growth  
RT factor-like domains: implications for the Marfan syndrome and other  
RT genetic disorders";  
RN Cell 85:597-605(1996).  
[9]  
RN REVIEW ON MFS VARIANTS.  
RP MEDLINE=96174615; PubMed=8594563;  
RX Colloid G., Beroud C., Soussi T., Junien C., Boileau C.;

RT factor-like domains: implications for the Marfan syndrome and other  
 RT genetic disorders";  
 RL Cell 85:597-605(1996).  
 RN [9].  
 RP REVIEW ON MFS VARIANTS.  
 RX MEDLINE=96174615; PubMed=8594563;  
 RA Colloid G., Beroud C., Soussi T., Junien C., Boileau C.;

"Software and database for the analysis of mutations in the human FNBI gene.";

RL Nucleic Acids Res. 24:137-141(1996).

[10]

REVIEW ON MFS VARIANTS.

RL MEDLINE=97169383; PubMed=9016526;

RL Colloid-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,

RL Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,

RL Richards R.I., Wang W., Junien C., Boileau C.;

RT "Marfan Database (second edition): software and database for the

RT analysis of mutations in the human FNBI gene.";

RL Nucleic Acids Res. 25:147-150(1997).

[11]

REVIEW ON VARIANTS.

RL MEDLINE=98062175; PubMed=9401003;

RL Hayward C., Brock D.J.H.;

RT "Fibrillin-1 mutations in Marfan syndrome and other type-1

RT fibrillinopathies.";

RL Hum. Mutat. 10:415-423(1997).

[12]

VARIANT MFS PRO-1137.

RL MEDLINE=91304569; PubMed=1852208;

RL Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,

RL Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,

RL Curristin S.M., Stetten G., Meyers D.A., Francomano C.A.;

RT "Marfan syndrome caused by a recurrent de novo missense mutation in

RT the fibrillin gene.";

RL Nature 352:337-339(1991).

[13]

VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.

RL MEDLINE=93250834; PubMed=1301946;

RL Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Francomano C.A.;

RT "Clustering of fibrillin (FNBI) missense mutations in Marfan syndrome

RT patients at cysteine residues in EGF-like domains.";

RL Hum. Mutat. 1:366-374(1992).

[14]

VARIANT MFS SER-2307.

RL MEDLINE=92235290; PubMed=1569206;

RL Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,

RL Corson G.M., Maslen C.L., Sakai L.Y., Francomano C.A., Cutting G.R.;

RT "Marfan phenotype variability in a family segregating a missense

RT mutation in the epidermal growth factor-like motif of the fibrillin

RT gene.";

RL J. Clin. Invest. 89:1674-1680(1992).

[15]

VARIANTS MFS ILE-548 AND ALA-723.

RL MEDLINE=94010946; PubMed=8406497;

RL Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,

RL Pyeritz R.E., Francomano C.A.;

RT "Four novel FNBI mutations: significance for mutant transcript level

RT and EGF-like domain calcium binding in the pathogenesis of Marfan

RT syndrome.";

RL Genomics 17:468-475(1993).

[16]

VARIANT MFS SER-2144.

RL MEDLINE=93278402; PubMed=8504310;

RL Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;

RT "A novel fibrillin mutation in the Marfan syndrome which could

RT disrupt calcium binding of the epidermal growth factor-like module.";

RL Hum. Mol. Genet. 2:475-477(1993).

[17]

VARIANTS MFS ARG-862; TYR-1117; PRO-1137 AND PHE-1589, AND VARIANT

RL ALA-1148.

RL MEDLINE=94108431; PubMed=8281141;

RL Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,

RL Berg M.A., Miller D.C., Francke U.;

RT "Mutation screening of complete fibrillin-1 coding sequence: report

RT of five new mutations, including two in 8-cysteine domains.";

RL Hum. Mol. Genet. 2:1813-1821(1993).

[18]

VARIANTS MFS GLY-217 AND ARG-2627.

RL MEDLINE=95067970; PubMed=7977366;

RL Karttunen L., Raghunath M., Loennekvist L., Peltonen L.;

[illegible]

Db 1420 CLNAPGGYRCBDMGFVPSADGKACEDIDCSLPNICVFGTCHNLPGLFRCECBIGYELD 1479  
 QY 456 RHIG-----TDCDSKGVGDGSGSEPPSPPTGSS-----TLTPPAYCLV 495  
 Db 1480 RSGNCTVNECLDPTTCISGCVN-----TPGSYICDPPDFELNTRVGCV 1527

RESULT 13  
 FBNI\_MOUSE  
 ID\_FBNI\_MOUSE STANDARD; PRT; 2871 AA.  
 AC Q61554; Q08026;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBNI OR FBNI-1.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP MEDLINE=95130561; PubMed=7829516;  
 RA Yin W., Germiller J., Sanguinetti C., Smiley E., Pangilinan T.,  
 RA Pereira L., Ramirez F., Bonadio J.;  
 RT "Primary structure and developmental expression of Fbn-1, the mouse  
 fibrillin gene.";  
 RL J. Biol. Chem. 270:1798-1806(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Kidney;  
 RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Structural component of connective tissue microfibrils  
 CC that binds calcium. Fibrillin-1-containing microfibrils provide  
 CC long-term force bearing structural support.  
 CC -!- PTM: Forms intramolecular disulfide bonds either with other  
 CC fibrillin-1 molecules or with other components of the microfibrils  
 CC (by similarity).  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBRP) domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>  
 CC or send an email to [license@iesb-sib.ch](mailto:license@iesb-sib.ch)).  
 CC -----  
 DR EMBL; L29454; AAA56840.1; -;  
 DR EMBL; U22493; AAA64217.1; -;  
 DR PIR; A55624; A55624.  
 DR HSSP; P35555; IAPJ.  
 DR MGD; MGI:95489; Fbni.  
 DR InterPro; IPR000152; Asx hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_CA.  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR002212; Fibril-assoc.  
 DR Pfam; PF00008; EGF; 45.  
 DR Pfam; PF00683; TB; 9.  
 DR SMART; SM00179; EGF\_CA; 42.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 43.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 38.  
 DR PROSITE; PS50026; EGF\_3; 45.  
 DR PROSITE; PS01187; EGF\_CA; 43.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 2871 FIBRILLIN 1.  
 FT DOMAIN 81 112 EGF-LIKE 1.  
 FT DOMAIN 115 146 EGF-LIKE 2.

147 DOMAIN  
 148 EGF-LIKE 3. CALCIIUM-BINDING.  
 149 EGF-LIKE 4. CALCIIUM-BINDING.  
 150 EGF-LIKE 5. CALCIIUM-BINDING.  
 151 PRO-RICH.  
 152 TGFBRP 1.  
 153 EGF-LIKE 6. CALCIIUM-BINDING.  
 154 EGF-LIKE 7. CALCIIUM-BINDING.  
 155 EGF-LIKE 8. CALCIIUM-BINDING.  
 156 EGF-LIKE 9. CALCIIUM-BINDING.  
 157 EGF-LIKE 10. CALCIIUM-BINDING.  
 158 TGFBRP 2.  
 159 EGF-LIKE 11. CALCIIUM-BINDING.  
 160 EGF-LIKE 12. CALCIIUM-BINDING.  
 161 EGF-LIKE 13. CALCIIUM-BINDING.  
 162 EGF-LIKE 14. CALCIIUM-BINDING.  
 163 TGFBRP 3.  
 164 EGF-LIKE 15. CALCIIUM-BINDING.  
 165 EGF-LIKE 16. CALCIIUM-BINDING.  
 166 EGF-LIKE 17. CALCIIUM-BINDING.  
 167 EGF-LIKE 18. CALCIIUM-BINDING.  
 168 EGF-LIKE 19. CALCIIUM-BINDING.  
 169 EGF-LIKE 20. CALCIIUM-BINDING.  
 170 EGF-LIKE 21. CALCIIUM-BINDING.  
 171 EGF-LIKE 22. CALCIIUM-BINDING.  
 172 EGF-LIKE 23. CALCIIUM-BINDING.  
 173 EGF-LIKE 24. CALCIIUM-BINDING.  
 174 EGF-LIKE 25. CALCIIUM-BINDING.  
 175 EGF-LIKE 26. CALCIIUM-BINDING.  
 176 TGFBRP 4.  
 177 EGF-LIKE 27. CALCIIUM-BINDING.  
 178 EGF-LIKE 28. CALCIIUM-BINDING.  
 179 TGFBRP 5.  
 180 EGF-LIKE 29. CALCIIUM-BINDING.  
 181 EGF-LIKE 30. CALCIIUM-BINDING.  
 182 EGF-LIKE 31. CALCIIUM-BINDING.  
 183 EGF-LIKE 32. CALCIIUM-BINDING.  
 184 EGF-LIKE 33. CALCIIUM-BINDING.  
 185 EGF-LIKE 34. CALCIIUM-BINDING.  
 186 EGF-LIKE 35. CALCIIUM-BINDING.  
 187 TGFBRP 6.  
 188 EGF-LIKE 36. CALCIIUM-BINDING.  
 189 EGF-LIKE 37. CALCIIUM-BINDING.  
 190 EGF-LIKE 38. CALCIIUM-BINDING.  
 191 EGF-LIKE 39. CALCIIUM-BINDING.  
 192 EGF-LIKE 40. CALCIIUM-BINDING.  
 193 TGFBRP 7.  
 194 EGF-LIKE 41. CALCIIUM-BINDING.  
 195 EGF-LIKE 42. CALCIIUM-BINDING.  
 196 EGF-LIKE 43. CALCIIUM-BINDING.  
 197 EGF-LIKE 44. CALCIIUM-BINDING.  
 198 EGF-LIKE 45. CALCIIUM-BINDING.  
 199 EGF-LIKE 46. CALCIIUM-BINDING.  
 200 EGF-LIKE 47. CALCIIUM-BINDING.  
 201 BY SIMILARITY.  
 202 BY SIMILARITY.  
 203 BY SIMILARITY.  
 204 BY SIMILARITY.  
 205 BY SIMILARITY.  
 206 BY SIMILARITY.  
 207 BY SIMILARITY.  
 208 BY SIMILARITY.  
 209 BY SIMILARITY.  
 210 BY SIMILARITY.  
 211 BY SIMILARITY.  
 212 BY SIMILARITY.  
 213 BY SIMILARITY.  
 214 BY SIMILARITY.  
 215 BY SIMILARITY.  
 216 BY SIMILARITY.  
 217 BY SIMILARITY.  
 218 BY SIMILARITY.  
 219 BY SIMILARITY.  
 220 BY SIMILARITY.  
 221 BY SIMILARITY.  
 222 BY SIMILARITY.  
 223 BY SIMILARITY.  
 224 BY SIMILARITY.  
 225 BY SIMILARITY.  
 226 BY SIMILARITY.  
 227 BY SIMILARITY.  
 228 BY SIMILARITY.  
 229 BY SIMILARITY.  
 230 BY SIMILARITY.  
 231 BY SIMILARITY.  
 232 BY SIMILARITY.  
 233 BY SIMILARITY.  
 234 BY SIMILARITY.  
 235 BY SIMILARITY.  
 236 BY SIMILARITY.  
 237 BY SIMILARITY.  
 238 BY SIMILARITY.  
 239 BY SIMILARITY.  
 240 BY SIMILARITY.  
 241 BY SIMILARITY.  
 242 BY SIMILARITY.  
 243 BY SIMILARITY.  
 244 BY SIMILARITY.  
 245 BY SIMILARITY.  
 246 BY SIMILARITY.  
 247 BY SIMILARITY.  
 248 BY SIMILARITY.  
 249 BY SIMILARITY.  
 250 BY SIMILARITY.  
 251 BY SIMILARITY.  
 252 BY SIMILARITY.  
 253 BY SIMILARITY.  
 254 BY SIMILARITY.  
 255 BY SIMILARITY.  
 256 BY SIMILARITY.  
 257 BY SIMILARITY.  
 258 BY SIMILARITY.  
 259 BY SIMILARITY.  
 260 BY SIMILARITY.  
 261 BY SIMILARITY.  
 262 BY SIMILARITY.  
 263 BY SIMILARITY.  
 264 BY SIMILARITY.  
 265 BY SIMILARITY.  
 266 BY SIMILARITY.  
 267 BY SIMILARITY.  
 268 BY SIMILARITY.  
 269 BY SIMILARITY.  
 270 BY SIMILARITY.  
 271 BY SIMILARITY.  
 272 BY SIMILARITY.  
 273 BY SIMILARITY.  
 274 BY SIMILARITY.  
 275 BY SIMILARITY.  
 276 BY SIMILARITY.  
 277 BY SIMILARITY.  
 278 BY SIMILARITY.  
 279 BY SIMILARITY.  
 280 BY SIMILARITY.  
 281 BY SIMILARITY.  
 282 BY SIMILARITY.  
 283 BY SIMILARITY.  
 284 BY SIMILARITY.  
 285 BY SIMILARITY.  
 286 BY SIMILARITY.  
 287 BY SIMILARITY.  
 288 BY SIMILARITY.  
 289 BY SIMILARITY.  
 290 BY SIMILARITY.  
 291 BY SIMILARITY.  
 292 BY SIMILARITY.  
 293 BY SIMILARITY.  
 294 BY SIMILARITY.  
 295 BY SIMILARITY.  
 296 BY SIMILARITY.  
 297 BY SIMILARITY.  
 298 BY SIMILARITY.  
 299 BY SIMILARITY.  
 300 BY SIMILARITY.  
 301 BY SIMILARITY.  
 302 BY SIMILARITY.  
 303 BY SIMILARITY.  
 304 BY SIMILARITY.  
 305 BY SIMILARITY.  
 306 BY SIMILARITY.  
 307 BY SIMILARITY.  
 308 BY SIMILARITY.  
 309 BY SIMILARITY.  
 310 BY SIMILARITY.  
 311 BY SIMILARITY.  
 312 BY SIMILARITY.  
 313 BY SIMILARITY.  
 314 BY SIMILARITY.  
 315 BY SIMILARITY.  
 316 BY SIMILARITY.  
 317 BY SIMILARITY.  
 318 BY SIMILARITY.  
 319 BY SIMILARITY.  
 320 BY SIMILARITY.  
 321 BY SIMILARITY.  
 322 BY SIMILARITY.  
 323 BY SIMILARITY.  
 324 BY SIMILARITY.  
 325 BY SIMILARITY.  
 326 BY SIMILARITY.  
 327 BY SIMILARITY.  
 328 BY SIMILARITY.  
 329 BY SIMILARITY.  
 330 BY SIMILARITY.  
 331 BY SIMILARITY.  
 332 BY SIMILARITY.  
 333 BY SIMILARITY.  
 334 BY SIMILARITY.  
 335 BY SIMILARITY.  
 336 BY SIMILARITY.  
 337 BY SIMILARITY.  
 338 BY SIMILARITY.  
 339 BY SIMILARITY.  
 340 BY SIMILARITY.  
 341 BY SIMILARITY.  
 342 BY SIMILARITY.  
 343 BY SIMILARITY.  
 344 BY SIMILARITY.  
 345 BY SIMILARITY.  
 346 BY SIMILARITY.  
 347 BY SIMILARITY.  
 348 BY SIMILARITY.  
 349 BY SIMILARITY.  
 350 BY SIMILARITY.  
 351 BY SIMILARITY.  
 352 BY SIMILARITY.  
 353 BY SIMILARITY.  
 354 BY SIMILARITY.  
 355 BY SIMILARITY.  
 356 BY SIMILARITY.  
 357 BY SIMILARITY.  
 358 BY SIMILARITY.  
 359 BY SIMILARITY.  
 360 BY SIMILARITY.  
 361 BY SIMILARITY.  
 362 BY SIMILARITY.  
 363 BY SIMILARITY.  
 364 BY SIMILARITY.  
 365 BY SIMILARITY.  
 366 BY SIMILARITY.  
 367 BY SIMILARITY.  
 368 BY SIMILARITY.  
 369 BY SIMILARITY.  
 370 BY SIMILARITY.  
 371 BY SIMILARITY.  
 372 BY SIMILARITY.  
 373 BY SIMILARITY.  
 374 BY SIMILARITY.  
 375 BY SIMILARITY.  
 376 BY SIMILARITY.  
 377 BY SIMILARITY.  
 378 BY SIMILARITY.  
 379 BY SIMILARITY.  
 380 BY SIMILARITY.  
 381 BY SIMILARITY.  
 382 BY SIMILARITY.  
 383 BY SIMILARITY.  
 384 BY SIMILARITY.  
 385 BY SIMILARITY.  
 386 BY SIMILARITY.  
 387 BY SIMILARITY.  
 388 BY SIMILARITY.  
 389 BY SIMILARITY.  
 390 BY SIMILARITY.  
 391 BY SIMILARITY.  
 392 BY SIMILARITY.  
 393 BY SIMILARITY.  
 394 BY SIMILARITY.  
 395 BY SIMILARITY.  
 396 BY SIMILARITY.  
 397 BY SIMILARITY.  
 398 BY SIMILARITY.  
 399 BY SIMILARITY.  
 400 BY SIMILARITY.  
 401 BY SIMILARITY.  
 402 BY SIMILARITY.  
 403 BY SIMILARITY.  
 404 BY SIMILARITY.  
 405 BY SIMILARITY.  
 406 BY SIMILARITY.  
 407 BY SIMILARITY.  
 408 BY SIMILARITY.  
 409 BY SIMILARITY.  
 410 BY SIMILARITY.  
 411 BY SIMILARITY.  
 412 BY SIMILARITY.  
 413 BY SIMILARITY.  
 414 BY SIMILARITY.  
 415 BY SIMILARITY.  
 416 BY SIMILARITY.  
 417 BY SIMILARITY.  
 418 BY SIMILARITY.  
 419 BY SIMILARITY.  
 420 BY SIMILARITY.  
 421 BY SIMILARITY.  
 422 BY SIMILARITY.  
 423 BY SIMILARITY.  
 424 BY SIMILARITY.  
 425 BY SIMILARITY.  
 426 BY SIMILARITY.  
 427 BY SIMILARITY.  
 428 BY SIMILARITY.  
 429 BY SIMILARITY.  
 430 BY SIMILARITY.  
 431 BY SIMILARITY.  
 432 BY SIMILARITY.  
 433 BY SIMILARITY.  
 434 BY SIMILARITY.  
 435 BY SIMILARITY.  
 436 BY SIMILARITY.  
 437 BY SIMILARITY.  
 438 BY SIMILARITY.  
 439 BY SIMILARITY.  
 440 BY SIMILARITY.  
 441 BY SIMILARITY.  
 442 BY SIMILARITY.  
 443 BY SIMILARITY.  
 444 BY SIMILARITY.  
 445 BY SIMILARITY.  
 446 BY SIMILARITY.  
 447 BY SIMILARITY.  
 448 BY SIMILARITY.  
 449 BY SIMILARITY.  
 450 BY SIMILARITY.  
 451 BY SIMILARITY.  
 452 BY SIMILARITY.  
 453 BY SIMILARITY.  
 454 BY SIMILARITY.  
 455 BY SIMILARITY.  
 456 BY SIMILARITY.  
 457 BY SIMILARITY.  
 458 BY SIMILARITY.  
 459 BY SIMILARITY.  
 460 BY SIMILARITY.  
 461 BY SIMILARITY.  
 462 BY SIMILARITY.  
 463 BY SIMILARITY.  
 464 BY SIMILARITY.  
 465 BY SIMILARITY.  
 466 BY SIMILARITY.  
 467 BY SIMILARITY.  
 468 BY SIMILARITY.  
 469 BY SIMILARITY.  
 470 BY SIMILARITY.  
 471 BY SIMILARITY.  
 472 BY SIMILARITY.  
 473 BY SIMILARITY.  
 474 BY SIMILARITY.  
 475 BY SIMILARITY.  
 476 BY SIMILARITY.  
 477 BY SIMILARITY.  
 478 BY SIMILARITY.  
 479 BY SIMILARITY.  
 480 BY SIMILARITY.  
 481 BY SIMILARITY.  
 482 BY SIMILARITY.  
 483 BY SIMILARITY.  
 484 BY SIMILARITY.  
 485 BY SIMILARITY.  
 486 BY SIMILARITY.  
 487 BY SIMILARITY.  
 488 BY SIMILARITY.  
 489 BY SIMILARITY.  
 490 BY SIMILARITY.  
 491 BY SIMILARITY.  
 492 BY SIMILARITY.  
 493 BY SIMILARITY.  
 494 BY SIMILARITY.  
 495 BY SIMILARITY.  
 496 BY SIMILARITY.  
 497 BY SIMILARITY.  
 498 BY SIMILARITY.  
 499 BY SIMILARITY.  
 500 BY SIMILARITY.  
 501 BY SIMILARITY.  
 502 BY SIMILARITY.  
 503 BY SIMILARITY.  
 504 BY SIMILARITY.  
 505 BY SIMILARITY.  
 506 BY SIMILARITY.  
 507 BY SIMILARITY.  
 508 BY SIMILARITY.  
 509 BY SIMILARITY.  
 510 BY SIMILARITY.  
 511 BY SIMILARITY.  
 512 BY SIMILARITY.  
 513 BY SIMILARITY.









[illegible]

|          | N-LINKED      | (GLCNAC. . .) | (POTENTIAL). |
|----------|---------------|---------------|--------------|
| N-LINKED | (GLCNAC. . .) | (POTENTIAL).  |              |



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 9, 2004, 08:43:13 ; Search time 37.5 seconds  
(without alignments)  
4190.082 Million cell updates/sec

Title: US-09-509-994-2\_COPY\_19\_516

Perfect score: 2830  
Sequence: 1 APAEPQPGSQCVHDCFAL.....PSPTPGSTLTPPAVLGHVSG 498

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID        | Description        |
|------------|--------|-------------|--------|-----------|--------------------|
| 1          | 2826   | 99.9        | 625    | 4 Q8IV29  | Q8IV29 homo sapien |
| 2          | 2658   | 93.9        | 468    | 4 Q9UC32  | Q9UC32 homo sapien |
| 3          | 1987.5 | 70.2        | 580    | 6 Q8HZ48  | Q8HZ48 oryctolagus |
| 4          | 1790   | 63.3        | 577    | 11 Q35370 | Q35370 rattus norv |
| 5          | 1780.5 | 62.9        | 491    | 11 Q8BUB5 | Q8BUB5 mus musculu |
| 6          | 1400.5 | 49.5        | 461    | 11 P97883 | P97883 rattus norv |
| 7          | 578    | 20.4        | 652    | 4 Q8IXK1  | Q8IXK1 homo sapien |
| 8          | 578    | 20.4        | 757    | 4 Q9HCU0  | Q9HCU0 homo sapien |
| 9          | 566.5  | 20.0        | 765    | 11 Q91ZV1 | Q91ZV1 mus musculu |
| 10         | 566.5  | 20.0        | 765    | 11 Q91V98 | Q91V98 mus musculu |
| 11         | 431    | 15.2        | 1664   | 5 Q9TVQ2  | Q9TVQ2 caenorhabdi |
| 12         | 414    | 14.6        | 1574   | 11 Q88281 | Q88281 rattus norv |
| 13         | 386    | 13.6        | 1409   | 5 Q9VS89  | Q9VS89 drosophila  |
| 14         | 376    | 13.3        | 708    | 13 P87363 | P87363 gallus gall |
| 15         | 375    | 13.3        | 2906   | 11 Q9WUH9 | Q9WUH9 rattus norv |
| 16         | 369    | 13.0        | 1231   | 4 Q8IU11  | Q8IU11 homo sapien |

|    |       |      |      |    |        |                    |
|----|-------|------|------|----|--------|--------------------|
| 17 | 369   | 13.0 | 1246 | 4  | O75095 | 075095 homo sapien |
| 18 | 367.5 | 13.0 | 1184 | 4  | O86V58 | O86V58 homo sapien |
| 19 | 367   | 13.0 | 3857 | 11 | O88840 | O88840 mus musculu |
| 20 | 365   | 12.9 | 2809 | 4  | Q96JP8 | Q96JP8 homo sapien |
| 21 | 365   | 12.9 | 2809 | 4  | Q86SJ5 | Q86SJ5 homo sapien |
| 22 | 361   | 12.8 | 2872 | 11 | Q9WUH8 | Q9WUH8 rattus norv |
| 23 | 359   | 12.7 | 528  | 11 | Q9CXH8 | Q9CXH8 mus musculu |
| 24 | 358.5 | 12.7 | 893  | 6  | Q8MJK0 | Q8MJK0 cercopithec |
| 25 | 358.5 | 12.7 | 1174 | 11 | Q99K58 | Q99K58 mus musculu |
| 26 | 357   | 12.6 | 715  | 4  | Q7Z5X1 | Q7Z5X1 homo sapien |
| 27 | 357   | 12.6 | 937  | 4  | Q96FT5 | Q96FT5 homo sapien |
| 28 | 357   | 12.6 | 1016 | 4  | Q8NDE6 | Q8NDE6 homo sapien |
| 29 | 356   | 12.6 | 590  | 11 | Q8C088 | Q8C088 mus musculu |
| 30 | 356   | 12.6 | 1231 | 4  | Q8IU10 | Q8IU10 homo sapien |
| 31 | 353.5 | 12.5 | 746  | 4  | Q96HB9 | Q96HB9 homo sapien |
| 32 | 353.5 | 12.5 | 1256 | 4  | Q9NS15 | Q9NS15 homo sapien |
| 33 | 353.5 | 12.5 | 1382 | 4  | Q9H7K2 | Q9H7K2 homo sapien |
| 34 | 352.5 | 12.5 | 741  | 4  | Q96K89 | Q96K89 homo sapien |
| 35 | 351.5 | 12.4 | 1511 | 4  | O75412 | O75412 homo sapien |
| 36 | 351.5 | 12.4 | 1587 | 4  | O00508 | O00508 homo sapien |
| 37 | 351   | 12.4 | 1666 | 11 | Q8K4G1 | Q8K4G1 mus musculu |
| 38 | 350   | 12.4 | 1600 | 11 | Q8K4G0 | Q8K4G0 mus musculu |
| 39 | 347.5 | 12.3 | 589  | 13 | O7ZX63 | O7ZX63 xenopus lae |
| 40 | 346   | 12.2 | 729  | 11 | Q8BNE3 | Q8BNE3 mus musculu |
| 41 | 346   | 12.2 | 787  | 11 | Q8K061 | Q8K061 mus musculu |
| 42 | 342   | 12.1 | 5636 | 4  | Q96RW7 | Q96RW7 homo sapien |
| 43 | 341   | 12.0 | 576  | 4  | Q9Y3V7 | Q9Y3V7 homo sapien |
| 44 | 340   | 12.0 | 1062 | 11 | Q60789 | Q60789 mus musculu |
| 45 | 338.5 | 12.0 | 1289 | 5  | Q86AC3 | Q86AC3 dictyosteli |

## ALIGNMENTS

RESULT 1

Q8IV29 PRELIMINARY; PRT; 625 AA.

ID Q8IV29; TISSUE=Brain;

AC Q8IV29; Strausberg R;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Thrombomodulin (Fragment).

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_taxID=9606;

[1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R;

EL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC035602; AAH35602.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0005529; F:sugar binding; IEA.

DR GO; GO:0004888; F:transmembrane receptor activity; IEA.

DR GO; GO:0007596; P:blood coagulation; IEA.

DR InterPro; IPR000152; Asx hydroxylase.

DR InterPro; IPR001881; EGF Ca.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR001304; Lectin C.

DR InterPro; IPR001491; Thrombomodulin.

DR Pfam; PF00008; EGF; 3.

DR Pfam; PF00059; Lectin c; 1.

DR PRINTS; PR00907; THROMBOMODULN.

DR SMART; SM00034; CLECT; 1.

DR SMART; SM00181; EGF; 6.

DR SMART; SM00179; EGF; CA; 4.

DR PROSITE; PS00010; ASX HYDROXYL; 2.

DR PROSITE; PS00041; C TYPE LECTIN 2; 1.

DR PROSITE; PS01186; EGF 2; 2.

DR PROSITE; PS01187; EGF\_CA; 2.

FT NON\_TER 1 1  
SQ SEQUENCE 625 AA; 65876 MW; 3E5E161183A854B9 CRC64;  
Query Match 99.9%; Score 2825; DB 4; Length 625;  
Best Local Similarity 99.8%; Pred. No. 1.2e-225;  
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
Db 69 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 128  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPVCV 120  
Db 129 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPVCV 188  
QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAVSITYGTTPA 180  
Db 189 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAVSITYGTTPA 248  
QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCCHEACNA 240  
Db 249 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCCHEACNA 308  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
Db 309 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 368  
QY 301 DOHRCEDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGSCVPEVDPFCFRANCEYQC 360  
Db 369 DOHRCEDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGSCVPEVDPFCFRANCEYQC 428  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGVIYLDGFI 420  
Db 429 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGVIYLDGFI 488  
QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALRHIGTDCDSKGVDDSGSGSEPPPS 480  
Db 489 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALRHIGTDCDSKGVDDSGSGSEPPPS 548  
QY 481 PTFGSLTTPPAVLVHSG 498  
Db 549 PTFGSLTTPPAVLVHSG 566  
RESULT 2  
Q9UC32 PRELIMINARY; PRT; 468 AA.  
AC Q9UC32;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombomodulin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93293792; PubMed=8390446;  
RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;  
RT "Urinary thrombomodulin, its isolation and characterization.";  
RL J. Biochem. 113:433-440 (1993).  
DR HSSP; P07204; 1ZAQ.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005529; F:sugar binding; IEA.  
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.  
DR GO; GO:0007596; F:blood coagulation; IEA.  
DR InterPro; IPR000152; ASX hydroxyl S.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR001304; Lectin\_C.

DR InterPro; IPR001491; Thrombomodulin.  
DR Pfam; PF00008; EGF; 3.  
DR PRINTS; PR00059; lectin c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 6.  
DR PROSITE; PS00010; ASX HYDROXYL; 2.  
DR PROSITE; PS00441; C-TYPE LECTIN 2; 1.  
DR PROSITE; PS01186; EGF 2; 2.  
DR PROSITE; PS01187; EGF CA; 2.  
KW EGF-like domain.  
SQ SEQUENCE 468 AA; 49444 MW; 4BFE8E98EFB6A40 CRC64;  
Query Match 93.9%; Score 2658; DB 4; Length 468;  
Best Local Similarity 99.8%; Pred. No. 6.9e-212;  
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
Db 1 APAEPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60  
QY 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPVCV 120  
Db 61 GGVGRRRLWIGLQPLPGCGDPKRLGRLGFQWVTGDNNTSYSRWARLDLNGAPLCPVCV 120  
QY 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAVSITYGTTPA 180  
Db 121 AVSAAEATVPSEPIWEEOQCEVKADGFLCEFHFPATCRPLAVEPGAAAVSITYGTTPA 180  
QY 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCCHEACNA 240  
Db 181 ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCCHEACNA 240  
QY 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
Db 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCTGYRLAA 300  
QY 301 DOHRCEDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGSCVPEVDPFCFRANCEYQC 360  
Db 301 DOHRCEDVDDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDGSCVPEVDPFCFRANCEYQC 360  
QY 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGVIYLDGFI 420  
Db 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPNTQASCECEGVIYLDGFI 420  
QY 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALRHIGTDCDSKGVDDSGSGSEPPPS 468  
Db 421 CTDIDECENGFGCSGVCHNLPGTFECICGPDLSALRHIGTDCDSKGVDDSGSGSEPPPS 468  
RESULT 3  
Q8HZ48 PRELIMINARY; PRT; 580 AA.  
AC Q8HZ48;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombomodulin precursor.  
GN THBD.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Deming C.B., Kim A.Y., Rade J.J.;  
RT "Cloning of rabbit thrombomodulin cDNA sequence.";  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY138902; AAN15931.2; -.  
DR PIR; A33308; A33308.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005529; F:sugar binding; IEA.

DR GO:0004888; F:transmembrane receptor activity; IEA.  
 DR GO:0007596; P:blood coagulation; IEA.  
 DR InterPro: IPR000152; Asx hydroxyl\_S.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR001304; Lectin C.  
 DR InterPro: IPR001491; Thrombomodulin.  
 DR Pfam: PF00008; EGF; 4.  
 DR PRINTS: PR00907; THROMBOMODULN.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 6.  
 DR SMART: SM00179; EGF\_Ca; 4.  
 DR PROSITE: PS00010; ASX HYDROXYL; 2.  
 DR PROSITE: PS00041; C TYPE LECTIN\_2; 1.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 DR PROSITE: PS01187; EGF\_Ca; 2.  
 KW Signal.  
 FT SIGNAL.  
 SQ SEQUENCE 580 AA; 60115 MW; 8A328FD2E141F42 CRC64;

Query Match 70.2%; Score 1987.5; DB 6; Length 580;  
 Best Local Similarity 71.6%; Pred. No. 3e-156;  
 Matches 363; Conservative 27; Mismatches 98; Indels 19; Gaps 8;  
 QY 3 AEPQPGSQCVHDCFPALYPGPATFLNASQICDGLRGLMTVRSSVAADVVISILLNGDGG 62  
 Db 23 AEPQPGSQCVHDCFPALYPGPATFLNASQICDGLRGLMTVRSSVAADVVISILLNGDGG 62  
 QY 63 VGRRLWIGLQPLPGCGDPRKRLGRLGFWQVTDNNNTSYRWARLDLNGAPLCPGCLVAV 122  
 Db 82 ---PRLWIGLQPLPGCGDPRKRLGRLGFWQVTDNNNTSYRWARLDLNGAPLCPGCLVAV 122  
 QY 123 SAAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPCGAAAAA---VSIITGTFP 179  
 Db 139 SAAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPCGAAAAA---VSIITGTFP 179

QY 180 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAAREAPGAWDCSVENGCGCEHACN 239  
 Db 199 GARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAAREAPGAWDCSVENGCGCEHACN 239  
 QY 240 AIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLA 299  
 Db 259 ASAGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLA 299  
 QY 300 ADQHRCEVDVDCILRPSQPCQCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQ 359  
 Db 319 ADGHRCEVDVDCILRPSQPCQCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQ 359  
 QY 360 COPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGTYLDDGF 419  
 Db 379 CLPLGQ-NYRCICAEAGFAPVDPDEPHRCQMFQNTACPADCDPNTQASCECPGTYLDDGF 419  
 QY 420 ICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALARHIGTDC-----DSKVTGG-DS 472  
 Db 438 LCADIDECN-GYCODECNLPGSYECICGPDPSALARHIGTDC-----DSKVTGG-DS 472  
 QY 473 GSGEPPTPTGPT-LTPPAVGLVHSG 498  
 Db 495 GSGEPPTPTGPT-LTPPAVGLVHSG 498

RESULT 4.  
 Q35370 ID O35370 PRELIMINARY; PRT; 577 AA.  
 AC O35370  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Thrombomodulin.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Lung;  
 RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-136 FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF022743; AAB80760.1; -  
 DR EMBL: AF022742; AAB80923.1; -  
 DR HSP: P07204; IFGD.  
 DR GO:0016020; C:membrane; IEA.  
 DR GO:0005509; F:calcium ion binding; IEA.  
 DR GO:0005529; F:sugar binding; IEA.  
 DR GO:0004888; F:transmembrane receptor activity; IEA.  
 DR GO:0007596; F:blood coagulation; IEA.  
 DR InterPro: IPR000152; Asx hydroxyl\_S.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR001304; Lectin C.  
 DR InterPro: IPR001491; Thrombomodulin.  
 DR Pfam: PF00008; EGF; 3.  
 DR Pfam: PF00059; Lectin C; 1.  
 DR PRINTS: PR00907; THROMBOMODULN.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00179; EGF\_Ca; 2.  
 DR PROSITE: PS00010; ASX HYDROXYL; 2.  
 DR PROSITE: PS00041; C TYPE LECTIN\_2; 1.  
 DR PROSITE: PS01186; EGF\_2; 3.  
 DR PROSITE: PS01187; EGF\_Ca; 2.  
 KW EGF-like domain.  
 SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;

Query Match 63.3%; Score 1790; DB 11; Length 577;  
 Best Local Similarity 63.9%; Pred. No. 6.8e-140;  
 Matches 321; Conservative 48; Mismatches 127; Indels 6; Gaps 5;  
 QY 1 APAEPQPGSQCVHDCFPALYPGPATFLNASQICDGLRGLMTVRSSVAADVVISILLNGD 60  
 Db 19 ALAKLQPKGSCQVNECFALFQDPVTFDASQORLQHLMTVRSSVAADVVISLLVS-D 77  
 QY 61 GVGRRRLWIGLQPLPGCGDPRKRLGRLGFWQVTDNNNTSYRWARLDLNGAPLCPGCLV 120  
 Db 78 SSMDSRP-WIGLQPLPGCGDPRKRLGRLGFWQVTDNNNTSYRWARLDLNGAPLCPGCLV 136  
 QY 121 AVSAAEAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEP-GAAAAVSIITGTFP 179  
 Db 137 TVSTATEAAGEPAWEKEPCENETKGFCEFFAAFCRPLRVNTRDPGAHISSTYNTPL 196  
 QY 180 AARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAAREAPGAWDCSVENGCGCEHACN 239  
 Db 197 GVGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAAREAPGAWDCSVENGCGCEHACN 256  
 QY 240 AIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLA 299  
 Db 257 RSANGPRCVCPGSGDLQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLA 316  
 QY 300 ADQHRCEVDVDCILRPSQPCQCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQ 359  
 Db 317 ADGHRCEVDVDCILRPSQPCQCVNTQGGFCHCYPNYDLVDGECVEPVDPCFRANCEYQ 376  
 QY 360 COPLNQTSLVCVCAEGFAPIPHEPHRCQMFQNTACPADCDPNTQASCECPGTYLDDGF 419  
 Db 377 CQVNSTHYNCICAEAGFAPVDPDEPHRCQMFQNTACPADCDPNTQASCECPGTYLDDGF 436  
 QY 420 ICTDIDECENGFGCSGVCHNLPGTFECICGPDPSALARHIGTDCSGKY--DGGDSGSGEP 477  
 Db 437 ICTDIDECQGECLTNECRNLPGSYECICGPDPSALARHIGTDCSGKY--DGGDSGSGEP 496  
 QY 478 PPS-PTFGSTLTTPPAVGLVHSG 498





```

Db 312 EGFILDRSGSICDIDECQCELTNECRNLPGSYECICGPTALAGQISKDCDPIPVLED 371
QY 469 GDSGSGSGPPPS-PTPGSTLTTPPAVLVHSG 498
Db 372 SEDGSGSGEFPSSNPTVVSSTVPSPSARDMHSG 402

RESULT 7
Q81XK1
ID Q81XK1 PRELIMINARY; PRT; 652 AA.
AC Q81XK1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Clq receptor protein precursor.
GN CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] _SEQUENCE FROM N.A.
RA Steinberger P., Stoeckl J., Wille S., Szekeres A., Prager E.,
RA Staffler G., Kuenig S., Kohl P., Majdic O., Knapp W., Stockinger H.;
RT "Identification of CD93 as the Clq receptor protein (ClqR) by
RL retroviral expression cloning.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ295142; CAC82720.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR InterPro; IPR000152; ASX_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001304; Lectin C.
DR InterPro; IPR001187; Tissue_factor.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00179; EGF_Ca; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
DR Signal; Receptor.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 652 C1Q RECEPTOR PROTEIN.
SQ SEQUENCE 652 AA; 69521 MW; 97549BA62CAF225E CRC64;

Query Match
Best Local Similarity 31.6%; Score 578; DB 4; Length 652;
Matches 175; Conservative 52; Mismatches 201; Indels 126; Gaps 31;

QY 6 QPGSGS-----CUEHDFCALYPGPATFLNASQICDGLRGLHMTVRSVAAD-----VI 53
Db 17 QPGAGTGADTEAVVVCVTACTYTAHSGKLSAAEAQNHCNQNGNLAVKSEEAQHVRVL 76
QY 54 SILLINGDGVGR-RLWIGLQLPPG-CDDPKGLPLRGFWVTGDNNTSYSRWARLDLN 110
Db 77 AQLLRREALTAMSKFWFIQLQREKQKCLDPSL--PLKGFVWVGGEDTFYSNWHKELRN 134
QY 111 GAPLCGLPLCA--VSAEAATVPSE-PINEEQQC-----EVKADGFLCFEHPFATCRPLA 161
Db 135 SC--ISKRCVSLLDLSQPLLSRLPKWSEGGCGSPGSGNIEGVCKFSFKGMCRLPLA 192
QY 162 V-EPGAAAANVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMC-TAPPGAVQGHV-- 217
Db 193 LGPG-----QVTVTPPTQTSSSLEAVPFAANVA-----CGEGDKDQTSQHYFL 239

```

```

QY 218 -AREAFGAWD-----CSVENGGCEHAC-NAIPGAPRCOCPAALQADGRS 261
Db 240 CEKADPVEDWGSSGFLCVSPKYGKFNNGGCHQDCFEFGDGSFLCGCRFRLDDLV 299
QY 262 CTASATQSCNDLCE--HFCVNPDPQGSYSYCMCETGYRLAADQHRCEDVDDCILEPSP 319
Db 300 C-ASRNPCSSPCRGATCALGP-HGKNYTCRCFQGYQLDSSQLDCVDVDEC--QSPCA 355
QY 320 QRCVNTQGGFECHCYENYDLVDEGCEVPEVDPCFRANCEYQCOPLNQTSLYCVCAEGFAP 379
Db 356 QECVNTPGGFRCECWGY-----EPGGP-----GEGACQDVDE-----CALGRSP- 395
QY 380 PHEPHRCQMFCNQATCAPADCDENTQAS--CECEGYIL--DDGFICTDIDEC--ENGGPC 433
Db 396 -----CAQGC-TNTDGSFHCSEGYVLAGEGTQCQDVDECVGFGPLC 439
QY 434 SGVCHNLPGTFCICGPDALSARHIGTDCDSKV-----DGGDSGSGE-----P 477
Db 440 DSLCFNTQGSFHCGLPGWLAPN-GVSCTMGPVSLGPPSPDDEDKGEKGSVTPRAA 498
QY 478 PPSPTPGSTLTTPPA 491
Db 499 TASPTRGEGTPKA 512

RESULT 8
Q9HCUO
ID Q9HCUO PRELIMINARY; PRT; 757 AA.
AC Q9HCUO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Tumor endothelial marker 1 precursor (Endosialin protein).
GN TM1 OR ENDOSIALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] _SEQUENCE FROM N.A.
RP MEDLINE=20407466; PubMed=10947988;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium.";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269274; PubMed=11084048;
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT "Molecular cloning and characterization of Endosialin, a C-type
RT lectin-like cell surface receptor of tumor Endothelium.";
RL J. Biol. Chem. 276:7408-7414(2001).
DR EMBL; AF275142; AAG00867.1; -.
DR EMBL; AJ295846; CAC34381.1; -.
DR HSSP; P07204; 1ZAQ.
DR Genew; HGNC:18219; CD164L1.
DR GO; GO:0005578; C:extracellular matrix; TAS.
DR InterPro; IPR000152; ASX_hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_Ca; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
KW EGF-like domain; Signal.

```

| FT   | SIGNAL  | 1   | 17        | POTENTIAL.                  |
|--|---|---|-----------|-----------------------------|
| ST   | CHAIN   | 18  | 757       | TUMOR ENDOTHELIAL MARKER 1. |
| SQ   | SEQUENCE  | 757 AA;   | 80858 MW; | C96363EALFD8FFA0 CRC64;     |
| Query Match 20.4%; Score 578; DB 4; Length 757;                    |   |   |           |                             |
| Best Local Similarity 33.6%; Pred. No. 2.2e-39;                    |   |   |           |                             |
| Matches 170; Conservative 41; Mismatches 181; Indels 114; Gaps 23; |   |   |           |                             |
| QY   | 3   | APQPQGGSCVHDCFALYFGPATFLNASQICDGLRGHLMTVRSVAAADVLSLLINGDGG    | 62        |                             |
| DB   | 23  | ASPR-----AACGSSCVAFPRRTTFLEAWRACRELGGLATPRPEEAQRVDSLVG--AG    | 77        |                             |
| QY   | 63  | VGRRLWIGLQLPFGCGDKPLRPLRGFWTGTDNNTSYSRMARLDLNGAPLCPGLCVAV     | 122       |                             |
| DB   | 78  | PASRLIWLIGLQARQCQLQR--PLRGFTWTGTDDQDTAFTNMAQ--PASGGPCPAQRCAVL | 134       |                             |
| QY   | 123   | SAAEATVPSEPIWEBQQCEVKADGFLCEPHFPATCRPLAVPFGAAAASVITYCTPPAAR   | 182       |                             |
| DB   | 135   | EAS-----GEHRWEGSCTIADVGLVCOFGPEGACPALQDEAGQGPV-----YTPPHLV    | 186       |                             |
| QY   | 183   | GADFOALVGGSAAY--APLGLQIMCTAPPGAVGQVHWAREAP---GAWDCSVENGCGEH   | 236       |                             |
| DB   | 187   | STEFNWLFPGSAAYVQCGAGRGASLLVKQPEGVG--WSRAGPLCLGT--GCPDNGGCGEH  | 244       |                             |
| QY   | 237   | AC-NAIPGAPRCQCPAGAALOAGRSCTASATQS--CNDLCEHFVCPNPDPQSGYSCMCET  | 294       |                             |
| DB   | 245   | BCVEEVDGHVSCRCCTEGFRLLADGRSCDDPCAQAPCEOOCE-----PGGPGQGYSCHRL  | 298       |                             |
| QY   | 295   | GYRLAADO-HRCEDVDDCILPEPSPQRCVNTQGGFECHCYENYDLDVDCGEVBPVDCFR   | 353       |                             |
| DB   | 299   | GFRPAEDDPHRCVDTDECOI-AGVCQCMVNVYGVGFEYCSEGELE-----            | 344       |                             |
| QY   | 354   | ANCEYQCPILMOTSYLVCVCAEGFAPLPEPHRCQFCMCTACPADCDP-----NTQASCEC  | 409       |                             |
| DB   | 345   | -----ADGIS-----   | 364       |                             |
| QY   | 410   | PEGYILDGFTCTIDE---CENGFGCSGVCHNLPGT--PECICGPDALSARHIGTDCDS    | 464       |                             |
| DB   | 365   | GD-BLLDDGEDEDEDAWKAFFNGWT-----EMPGLIWMETQPPDFALAYRSPED-       | 417       |                             |
| QY   | 465   | GKVDGSDSGSGEPSPPTGSGTLTPP                                     | 490       |                             |
| DB   | 418   | -----REFQIPYPTWPPPP   | 432       |                             |
| RESULT 9   |   |   |           |                             |
| Q91ZV1 PRELIMINARY; PRT; 765 AA.                                   |   |   |           |                             |
| ID   | Q91ZV1  |   |           |                             |
| AC   | Q91ZV1  |   |           |                             |
| DT   | 01-DEC-2001 (Tremblrel. 19, Created)                                  |   |           |                             |
| DT   | 01-DEC-2001 (Tremblrel. 19, Last sequence update)                     |   |           |                             |
| DT   | 01-OCT-2003 (Tremblrel. 25, Last annotation update)                   |   |           |                             |
| DE   | Endosomal.  |   |           |                             |
| GN   | TEM1.   |   |           |                             |
| OS   | Mus musculus (Mouse).   |   |           |                             |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |   |           |                             |
| OX   | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.    |   |           |                             |
| NCBI_TaxID=10090;  |   |   |           |                             |
| RN   | [1]   |   |           |                             |
| RP   | SEQUENCE FROM N.A.  |   |           |                             |
| RC   | STRAIN=Swiss Webster; TISSUE=Thymus;                                  |   |           |                             |
| RX   | MEDLINE=21486432; PubMed=1148985;                                     |   |           |                             |
| RA   | Opavsky R., Haviernik P., Jurkovicova D., Garin M.T., Copeland N.G.,  |   |           |                             |
| RA   | Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,     |   |           |                             |
| RA   | Oue A., Wolff L.;   |   |           |                             |
| RT   | "Molecular Characterization of the Mouse Tem1/endosomal Gene          |   |           |                             |
| RT   | Regulated by Cell Density in Vitro and Expressed in Normal Tissues in |   |           |                             |
| RT   | Vivo.";   |   |           |                             |
| RL   | J. Biol. Chem. 276:38795-38807(2001).                                 |   |           |                             |
| RL   | EMBL; AF388572; AAK84664.1; -   |   |           |                             |
| DR   | MGD; MGI:1917695; Tem1.   |   |           |                             |
| DR   | GO; GO:0005737; Cytoplasm; IDA.                                       |   |           |                             |
| DR   | Interferon; IPR000152; Asx hydroxyl S.                                |   |           |                             |

RT "Molecular Characterization of the Mouse Tem1/endosomal Gene  
RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in  
RT Vivo.";  
RL J. Biol. Chem. 276:38795-38807(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Olfactory epithelium;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Olfactory epithelium;  
RA Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF378758; AAL11995.1; -  
DR EMBL; AF388573; AAK8465.1; -  
DR EMBL; BC046318; AAH46318.1; -  
DR MGD; MGI:1917695; Tem1.  
DR GO; GO:0005737; C:cytoplasm; IDA.  
DR InterPro; IPR000152; ASX\_HYDROXYL\_S.  
DR InterPro; IPR001881; EGF\_CA.  
DR InterPro; IPR006209; EGF\_LIKE.  
DR InterPro; IPR001304; Lectin\_C.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00008; EGF\_2; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi; 1.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00179; EGF\_CA; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS00041; C-TYPE LECTIN\_2; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR EGF-like domain; Signal.  
FT SIGNAL 1 17  
FT CHAIN 18 765  
FT CHAIN 18 765 TUMOR ENDOTHELIAL MARKER 1.  
SQ SEQUENCE 765 AA; 81813 MW; 572A06CC15B8C8D8 CRC64;

Query Match 20.0%; Score 566.5; DB 11; Length 765;  
Best Local Similarity 38.9%; Pred. No. 2e-38;  
Matches 139; Conservative 37; Mismatches 146; Indels 35; Gaps 16;  
2 PAEPQCGSCVCHDCFALYEGRATFNASQICDLRGHLMTVRSSVAADVLSILLNGDG 61  
20 PWTPEPRAA-CGPPSSCALPFRRTFLFAMRACRELGNLATPTPEEAQRVDSL- 74  
62 GVGRRR-LMIGLQPPCGDCKELRGLRGFWVTGNNTSYSWARLDLNGALPGPLC 119  
75 GVGPANGLLMIGLQARQACQQR-PLRGIWTTGQDFTFNWAPATEG-PCPAQR 131  
120 VAVSAAEATVPSEPIWEEQCEVADGFLCEFHFPATCRFLAVEPGAAAAVSTTPTFF 179  
132 AALEAS-----GHRWTEGSCVTADVGYLCQFGPEGACPALPLEVGAGPAV---YTPPF 183

QY 180 AARGADFOALPYGVSSAAV---APLGLQLMCTAPPAGVQGHWAREAP--GAWDCSVENGCC 234  
DB 184 NLVSEFEWLPFGSVAAVQCCAGRGASLLICVKQPSGGVG-WSQTGLPCPTGCGPDNGGC 242  
QY 235 EHAC-NAIFGAPRCQCPAGAAALQADGRSCTASATQS-CNDLCEHFQVFPDPQPSYSCHC 292  
DB 243 EHECVVEVDGAVSCRCSEGFRLAADGSHCEDPCAQAPCQQCE-----PGSPQYSCHC 296  
QY 293 ETYGRLAAQ-QHRCEDVDCLLEPSPCPQRCVNTQGFCHCVPNYDL-VDGECVFP 347  
DB 297 RLGFRAEDDPHRCVDTDCQI-AGVCQCMVNYVGGFECYCSGHELEADGISCSP 352  
RESULT 11  
Q9TVQ2 PRELIMINARY; PRT; 1664 AA.  
ID Q9TVQ2;  
AC Q9TVQ2;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Y64G10A.7 protein.  
GN Y64G10A.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Mortimore B.J.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RX SEQUENCE FROM N.A.  
RX MEDLINE=98069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ainecough R.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL117206; CAB60454.1; -  
DR EMBL; AL110498; CAB60454.1; JOINED.  
DR EMBL; AL110498; CAB57911.1; -  
DR EMBL; AL117206; CAB57911.1; JOINED.  
DR HSP; P00736; IAPQ.  
DR WormPep; Y64G10A.7; CE24549.  
DR GO; GO:000509; F:calcium ion binding; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR00152; Asx\_hydroxyl\_S.  
DR InterPro; IPR001881; EGF\_CA.  
DR InterPro; IPR006209; EGF\_LIKE.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR Pfam; PF00008; EGF; 22.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR SMART; SM00179; EGF\_CA; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 4.  
DR PROSITE; PS00022; EGF\_1; 22.  
DR PROSITE; PS01186; EGF\_2; 24.  
DR PROSITE; PS01187; EGF\_CA; 3.  
DR EGF-like domain.  
KW EGF-like domain.  
SQ SEQUENCE 1664 AA; 179279 MW; A69F093B4C705832 CRC64;  
Query Match 15.2%; Score 431; DB 5; Length 1664;  
Best Local Similarity 29.7%; Pred. No. 8e-27;  
Matches 126; Conservative 40; Mismatches 178; Indels 80; Gaps 19;  
101 YSRWARLDLNGALPGPLCVANSAE-----ATVPSEPIWEEQCE--VKADGFLCEP 151  
56 YLFAFAPSRGSKCCLLRQVANCSDLCHNGGTCVPESEHNDNEQCEPCVGTGAKCQY 115  
152 HFPATCRFLAVEPGAAAAVSTTPTPEAARGADFOALPVGSSAAVAPLGLQLMCTAPP 211

```

Db 116 D-ANEC--MANNOCCEHCYN-TIGTY--CRCPWG 145
QY 212 AVQHWAREAPGWDSCVNGGCEHACNAIPGAPRCOCAGALOADGRSCTA-----264
Db 146 FELSGDNTCSDDICAVSNGGSDRCVNGPFGFRCDPDLVHADGRTCGSGFHENL 205
QY 265 ---SATQSC---NDLCEHFCVNPDPQSGYSKCMCTGYRLAADQHRCEVDVDDCILEPSPC 318
Db 206 ILIKKVTSCSTDNGGCEHEC-ENDSNGEFYRCRCRVGFKLSKNSKQCPVDFCDNKGCC 264
QY 319 PRCVNTQGGFECHCVNYDL-VGCEVFPDPCFRAN-CEYOCQPLNQTSLYLCVCAEFG 376
Db 265 QHCTNNHGAQCCQYFGFHLSDYDRSCVDIDECANNGCEHFCENVKGT-YRCKREGY 323
QY 377 APIPEPHRCOMP-----CNQATCAPADC--DPNTQASCECEGYIL-DDGFICTDIDEC- 427
Db 324 -QLGRDRTCEEMLGCGVNGGCGCHDQYDQPGHGVCKRCRNYILLANDQKJCHDINECH 382
QY 428 ENGFGSCGVCHNLPGTFECICGPDSSALARIHGT-----GKVDGDSG 473
Db 383 ENNGDCSQICVNLAGSVEQCKPGLMKDRKTCEDISECSSNNGGCEQICSNQEGGYMC 442
QY 474 SGEP 477
Db 443 SCEP 446

RESULT 12
ID 088281 PRELIMINARY; PRT; 1574 AA.
AC 088281;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MEGF6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98360089; PubMed=98303030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RL like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL; AB011532; BAA32462.1; -.
DR PIR; T13954; T13954.
DR HSP; P00736; 1APO.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000152; ASX hydroxylase.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 20.
DR PRINTS; PR000011; EGF_LAMININ.
DR SMART; SM00179; EGF_CA; 4.
DR PROSITE; PS00010; ASX HYDROXYL; 5.
DR PROSITE; PS00022; EGF_1; 23.
DR PROSITE; PS01186; EGF_2; 23.
DR PROSITE; PS01187; EGF_CA; 5.
DR KW EGF-like domain.
SQ SEQUENCE 1574 AA; 165445 MW; 2B48533D8F7F6E7 CRC64;

Query Match
Best Local Similarity 34.1%; DB 11; Length 1574;
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

QY 206 CTAPPG-AVOGHWAREAPGWDSCVNGGCEHACNAIPGAPRCOCAGALOADGRSCTA 264
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

```

Db 147 CRCPGYQLQGD-GKTCQDVDECRANHGQHRVNTPGSYLCECKPGFRLHDTGRCL- 204
QY 265 SATQSC---NDLCEHFCVP-----NPDPGYSYSCM-----291
Db 205 -AISCTLNGGCGHOCVQLTQHRCCQCPQYQLQEDGRRCVRRSPCAEANGGCMHICQ 263
QY 292 -----CETGYRLAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCVNYDL-VD 341
Db 264 EURLGAHCCGFGYQLAADRKTCEDVDICALGLAQCAHGLCNTQGSFKCVCHAGVELGAD 323
QY 342 G-EC-----VEPYDPCFRAN-CEYOCQPLNQTSLYLCVCAEFGAPIPHEPHRCOMFC-----390
Db 324 GRQCVRIEMEIIVNSCEANGGCGSHGSH-TSTGPLCTCPRGY-----ELDEQKTCIDID 377
QY 391 ---NOTACPADCDPNTQA--SCECEGYILD-DGFICTDIDECENG-GFSCGVCHNLPGT 443
Db 378 DCANSPPCCQAC-ANTPGYEGSCFAGYRLNTDGGCEDVDCEASGHGCEHCNLAGS 436
QY 444 FECIC 448
Db 437 FOCFC 441

RESULT 13
QYVS89
ID QYVS89 PRELIMINARY; PRT; 1409 AA.
AC QYVS89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG7526 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Balow R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

```

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celnik S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon P.C., Rogers Y.,  
 RA Banzon J., An H., Balow D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferriera S., Frise E., Galle R.P., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Numoo J.,  
 RA Pacaleb J., Paragass V., Park S., Patel S., Pfeiffer B.,  
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AE003558; AAF50538.2;  
 DR HSSP; P00736; IAP0.  
 DR FlyBase; FEGN0035798; CG7526.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0004888; F:transmembrane receptor activity; IEA.  
 DR GO; GO:0007596; P:blood coagulation; IEA.  
 DR InterPro; IPR000152; Asx hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR003410; HVALin.  
 DR InterPro; IPR006210; IEGL.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PF00008; EGF; 11.  
 DR Pfam; PF00084; sushi; 2.  
 DR PRINTS; PR00907; THRBOMODULN.  
 DR SMART; SM00032; CCP; 2.  
 DR SMART; SM00181; EGF; 13.  
 DR SMART; SM00179; EGF\_Ca; 14.  
 DR PROSITE; PS00010; ASX HYDROXYL; 8.  
 DR PROSITE; PS01186; EGF\_2; 9.  
 DR PROSITE; PS01187; EGF\_Ca; 10.  
 KW EGF-like domain.  
 SQ SEQUENCE 1409 AA; 154131 MW; F9A0DE1478AF354C CRC64;

Query Match 13.6%; Score 386; DB 5; Length 1409;  
 Best Local Similarity 32.1%; Pred. No. 3.5e-23;  
 Matches 99; Conservative 33; Mismatches 110; Indels 56; Gaps 18;

QY 206 CTAPPGAVQV-----HWAREPAGW-----DCSVNGGCEHACNAIPG 243

Db 563 CLCPFGYALGLDNHIVTSLNSSFITDSTSETPSAHTCLDIDECSLANGNSHFCQNEPG 622  
 QY 244 AFRQCQAGAAALQADGRSCTASATOSC---NDLCEHFVCPNPDPQSGYSYCMCETGYRLAA 300  
 Db 623 GFQCACPLGLYALSEDWRTCC-QDIDECILDSNGQCSQLCL---NQFGGFACACETGFELTP 677  
 QY 301 DQHRCEVDVDDCILERSPCPCQRCVNTQGGFECYNYDVLVDGE---CVPVDPCC---FRAN 355  
 Db 678 DQFGCADIDECSDYGNCSYDINLLGTHACACERGYELAKDKLSCLD-VDECAGLGG 736  
 QY 356 CEYQCOPLNOT-SYLCVCAEGFAPRPHRQCFMFCNQATACEA---DCDNTQASC----- 407  
 Db 737 CSHECC---INAKTFEGCGPLG---LINDDGR-----SCSPALVCGPPTQRSADGCA 785  
 QY 408 -RCPEGYLDDGFICTDIDEC---ENGCGSGVCHNLPCTFEICIGP-----DSALARHI 458  
 Db 786 IECNPGYTLGSDDKCVDDIDECQKQNGG-CSHRCSNTEGSKFCSCPPGYELDSQDKTCQDI 844  
 QY 459 GTDCDSGK 466  
 Db 845 -DECDQDK 851  
 RESULT 14  
 P87363 PRELIMINARY; PRT; 708 AA.  
 AC P87363;  
 DT 01-MAY-1997 (TREMELrel. 03, Created)  
 DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Fibrillin-1 (Fragment).  
 GN FBNI.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]\_TaxID=9031;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20152896; PubMed=10691037;  
 RA Zhou G., Price C.E., Rosenquist T.H., Gadson P.F., Godfrey M.;  
 RT "Partial cloning and sequencing of chick fibrillin-1 cDNA.";  
 RL In Vitro Cell. Dev. Biol. Anim. 36:19-25 (2000).  
 DR EMBL; U88872; AA848531.1; -;  
 DR HSSP; P07204; 2A2X.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR InterPro; IPR000152; Asx hydroxyl\_S.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR002212; Fibrin-assoc.  
 DR Pfam; PF00008; EGF; 13.  
 DR Pfam; PF00683; TB; 2.  
 DR SMART; SM00179; EGF\_Ca; 14.  
 DR PROSITE; PS00010; ASX HYDROXYL; 13.  
 DR PROSITE; PS01186; EGF\_2; 10.  
 DR PROSITE; PS01187; EGF\_Ca; 13.  
 KW EGF-like domain.  
 FT NON TER 708  
 SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

Query Match 13.3%; Score 376; DB 13; Length 708;  
 Best Local Similarity 31.8%; Pred. No. 1.1e-22;  
 Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

QY 226 DCSVNGGCEHACNAIPGRCQCPAGAAALQADGRSCTASATOSCNDLCEHFVCPNP--- 282  
 Db 223 ECSIMNGGCENFCTGSEGSYECSCKQGFALPDRTCT-----DDE--CEDNPNIC 272  
 QY 283 -----DQPGYSYCMCETGYRLAADQHRCEVDVDDCILPEPSPC-FQRCVNTQGGFECYCP 335  
 Db 273 DGGQCTNIPGEYRCLCYDGFVASEDMKTCVDVNECDLHFNICLSGTCTKSGSFICDCM 332

```
QY 336 NYDLVDGE--CVEPVDPC--FRANCEYQCQPLN-QTSYLCVCAEGFA-----PIP 380
Db 333 GYSKKGTTGCTD-INECEIGAHCNDRHAVCTNIPGSFKSCSSGWIENGKICTDLDECS 391
QY 381 HEPHRCQMFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDE-----CENG- 430
Db 392 NGTHKCSPH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDCESENLMNCENGQ 442
QY 431 -----GF-----CS-----GVCHNLPCTFECICGPDLSALA 455
Db 443 CLNAPGGYRCEDMGFLPSLDGKACEDIDEGSLNLCVYGTCHNLPGLFRCECEVGYELD 502
QY 456 RHIG-----TDCDSK-VDGSDSGSEPPSPFTFGSTLTPPAVLIV 495
Db 503 RSGGNCCTDVNECADPTTCISGTCVNTAGSYTCECPD-----FELNPTRVGCV 550

RESULT 15
Q9WUH9 PRELIMINARY; PRT; 2906 AA.
AC Q9WUH9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibrillin-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99350231; PubMed=10419698;
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E.,
RA Kanwar Y.S.;
RT "Cloning of rat fibrillin-2 cDNA and its role in branching
RT morphogenesis of embryonic lung.";
RL Dev. Biol. 212:229-242(1999).
DR EMBL; AF135060; AAD34439.1; -.
DR HSP; P3555; IEMN.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR000152; Aex_Hydroxyl_S.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 44.
DR Pfam; PF00583; TB; 9.
DR SMART; SM00179; EGF_CA; 42.
DR PROSITE; PS00070; ALDEHYDE DEHYDR_CYS; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
DR EGF-like domain.
KW EGF-like domain.
SQ SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;
```

```
Query Match 13.3%; Score 375; DB 11; Length 2906;
Best Local Similarity 32.9%; Pred. No. 6.5e-22;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 226 DCSVENGGCHACNAIPGAPROCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPD-- 283
Db 1237 ECMIMNGGCDTQCTNSRGSEYCSGYALMPDGRSCA-----DIDE--CENNPDI 1286
QY 284 -----QPGSYCMCTGYRLAADQRCEDVDCCILEPSPCP-QRCVNTQGGFECHCYP 335
Db 1287 DGGQCTNIPGEYRCLCVDFGFMASMDMTCTIDVNECDLNPNICMFGECNTKGSFICHQQL 1346
QY 336 NYDLVDGE--CVEPVDPC--FRANCEYQCQPLN-QTSYLCVCAEGFA-----PIP 380
Db 1347 GYSVKKGATGCTD-VDECEIGAHCNCDMHWASCLNVPGSFKSCREGWNGIKCIDLDECA 1405
```

```
QY 381 HEPHRCQMFNCQTACPADCDPNTQAS--CECPGYILDDGFICTDIDE-----CENG- 430
Db 1406 NGTHQCSI-----NAQC-VNTPGSYRCACSEGF-TGDGFTCSVDCEAENINLCENGQ 1456
QY 431 -----GF-----CS-----GVCHNLPCTFECICGPDLSALA 455
Db 1457 CLNVPGAYRCECEMGFTPASDSRSCQIDECSPQNICVFGTCNNLPGMFHCICDDGYGLD 1516
QY 456 RHIG--TDCC 463
Db 1517 RTGGHCTDID 1526
```

Search completed: June 9, 2004, 08:50:01  
Job time : 39.5 secs